

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)


```
; FILE REFERENCE: 200755/1002
; CURRENT APPLICATION NUMBER: US/09/146.580A
; CURRENT FILING DATE: 1998-09-03
; EARLIER APPLICATION NUMBER: 60/071.889
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 60/092.155
; EARLIER FILING DATE: 1998-07-09
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 104
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-146-580-8

Query Match          5.2%; Score 104; DB 4; Length 104;
Best Local Similarity 100.0%; Pred. No. 1.5e-35;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 745 actggattatagtgtagccacagtcctggcctgaattttctgtgatacaaatc 804
      |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 104 ACTGGATTATAGGTGTAGGCCACAGTCCTGGCCTAATATTTCTGTGATCAAAATC 45
      |||||||||||||||||||||||||||||||||||||||||||||||||||

Qy 805 aggttaatgttttggttaagaatttcctacgtgaattcgtgt 848
      |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 44 AGGTTAATGTTTGGTTAGAAATTCCTAGTGAATCGTGT 1

RESULT 5
US-09-033-333-3
; Sequence 3, Application US/09033333
; Patent No. 6197293
; GENERAL INFORMATION:
; APPLICANT: Yu, De Chao
; APPLICANT: Schuur, Eric
; APPLICANT: Henderson, Daniel
; TITLE OF INVENTION: ADENOVIRUS VECTORS SPECIFIC
; TITLE OF INVENTION: FOR CELLS EXPRESSING ANDROGEN RECEPTOR AND METHODS OF USE
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/033.333
; FILING DATE: 02-MAR-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Catherine, Polizzi M
; REGISTRATION NUMBER: 40,130
; REFERENCE/DOCKET NUMBER: 34802-20007.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-813-5600
; TELEFAX: 650-494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5835 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
```

```
; TOPOLOGY: linear
US-09-033-333-3

Query Match          2.6%; Score 52; DB 4; Length 5835;
Best Local Similarity 100.0%; Pred. No. 2.3e-13;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1169 tttttttgagacagagtcctcactctgtgtccacagctgagtgcaagtggc 1220
      |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3899 TTTTGTGAGACAGAGTCTCACTCTGTGTGCCAGGCTGGAGTGCAGTGCC 3950

RESULT 6
US-08-380-916-1
; Sequence 1, Application US/08380916
; Patent No. 5648478
; GENERAL INFORMATION:
; APPLICANT: Calydon, Inc.
; TITLE OF INVENTION: Tissue-Specific Enhancer Active in
; TITLE OF INVENTION: Prostrate
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/380.916
; FILING DATE: 12-JAN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/182,247
; FILING DATE: 13-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I.
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: EP-60058-PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5836 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-380-916-1

Query Match          2.6%; Score 52; DB 1; Length 5836;
Best Local Similarity 100.0%; Pred. No. 2.3e-13;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1169 tttttttgagacagagtcctcactctgtgtccacagctgagtgcaagtggc 1220
      |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3900 TTTTGTGAGACAGAGTCTCACTCTGTGTGCCAGGCTGGAGTGCAGTGCC 3951

RESULT 7
US-08-721-690-1
; Sequence 1, Application US/08721690
; Patent No. 6057299
; GENERAL INFORMATION:
; APPLICANT: Henderson, Daniel R.
```

;; TITLE OF INVENTION: TISSUE-SPECIFIC ENHANCER ACTIVE
;; TITLE OF INVENTION: IN PROSTATE
;; NUMBER OF SEQUENCES: 22
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: MORRISON & FOERSTER
;; STREET: 755 PAGE MILL ROAD
;; CITY: Palo Alto
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94304-1018
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: DOS
;; SOFTWARE: FastSeq for Windows Version 2.0
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/721,690
;; FILING DATE: 27-SEP-1996
;; CLASSIFICATION: 514
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/380,916
;; FILING DATE: 30-JAN-1995
;; APPLICATION NUMBER: US 08/182,247
;; FILING DATE: 13-JAN-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Catherine, Polizzi M
;; REGISTRATION NUMBER: 40,130
;; REFERENCE/DOCKET NUMBER: 34802-20001.21
;; TELEPHONE: 415-813-5600
;; TELEFAX: 415-494-0792
;; TELEX: 706141
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 5836 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; US-08-721-690-1

Query Match 2.6%; Score 52; DB 3; Length 5836;
Best Local Similarity 100.0%; Pred. No. 2.3e-13;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1169 ttttttttgagacagctcactctgttgcacagctggaagtgcagtggc 1220
|||||
Db 3900 TTTTGTGACAGAGTCTACTCTGTGCCCCAGGCTGGAGTGCAGTGGC 3951

RESULT 8
US-08-891-581-1
;; Sequence 1, Application US/08891581
;; Patent No. 6136792
;; GENERAL INFORMATION:
;; APPLICANT: Henderson, Daniel R.
;; TITLE OF INVENTION: TISSUE SPECIFIC ENHANCER ACTIVE
;; TITLE OF INVENTION: IN PROSTATE
;; NUMBER OF SEQUENCES: 2
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: MORRISON & FOERSTER
;; STREET: 755 PAGE MILL ROAD
;; CITY: Palo Alto
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94304-1018
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: DOS
;; SOFTWARE: FastSeq for Windows Version 2.0
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/891,581

;; FILING DATE:
;; CLASSIFICATION: 514
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/380,916
;; FILING DATE: 30-JAN-1995
;; APPLICATION NUMBER: US 08/182,247
;; FILING DATE: 13-JAN-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Catherine, Polizzi M
;; REGISTRATION NUMBER: 40,130
;; REFERENCE/DOCKET NUMBER: 34802-20001.22
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 415-813-5600
;; TELEFAX: 415-494-0792
;; TELEX: 706141
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 5836 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; US-08-891-581-1

Query Match 2.6%; Score 52; DB 3; Length 5836;
Best Local Similarity 100.0%; Pred. No. 2.3e-13;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1169 ttttttttgagacagctcactctgttgcacagctggaagtgcagtggc 1220
|||||
Db 3900 TTTTGTGACAGAGTCTACTCTGTGCCCCAGGCTGGAGTGCAGTGGC 3951

RESULT 9
US-09-033-333-2
;; Sequence 2, Application US/09033333
;; Patent No. 6197293
;; GENERAL INFORMATION:
;; APPLICANT: Yu, De Chao
;; APPLICANT: Schuur, Eric
;; APPLICANT: Henderson, Daniel
;; TITLE OF INVENTION: ADENOVIRUS VECTORS SPECIFIC
;; TITLE OF INVENTION: FOR CELLS EXPRESSING ANDROGEN RECEPTOR AND METHODS OF USE
;; TITLE OF INVENTION: THEREOF
;; NUMBER OF SEQUENCES: 22
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: MORRISON & FOERSTER
;; STREET: 755 PAGE MILL ROAD
;; CITY: Palo Alto
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94304-1018
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: Windows
;; SOFTWARE: FastSeq for Windows Version 2.0b
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/033,333
;; FILING DATE: 02-MAR-1998
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Catherine, Polizzi M
;; REGISTRATION NUMBER: 40,130
;; REFERENCE/DOCKET NUMBER: 34802-20007.00
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 650-813-5600
;; TELEFAX: 650-494-0792
;; TELEX: 706141
;; INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 5836 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-033-333-2

Query Match 2.6%; Score 52; DB 4; Length 5836;
Best Local Similarity 100.0%; Pred. No. 2.3e-13;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1169 ttttttttgagacagagctcactctgttgcacagctggagtgcaagtggc 1220
|||||
Db 3900 TTTTITTTTGACACAGAGTCTCAGTCTGTGTCACAGGCTGGAGTGCAGTGGC 3951

RESULT 10

US-09-167-681-45
Sequence 45, Application US/09167681A
Patent No. 6265561

GENERAL INFORMATION:

APPLICANT: Weinshilboum, M.D., Richard M.
APPLICANT: Raftogiannis, Rebecca B.
APPLICANT: Wood, Thomas C.
APPLICANT: Ottensmeyer, Diane M.

TITLE OF INVENTION: SULFOTRANSFERASE SEQUENCE VARIANTS

FILE REFERENCE: 07039/118001

CURRENT APPLICATION NUMBER: US/09/167,681A

CURRENT FILING DATE: 1998-10-07

NUMBER OF SEQ ID NOS: 52

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 45

LENGTH: 8447

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (4361)...(4507)

NAME/KEY: CDS

LOCATION: (4612)...(4737)

NAME/KEY: CDS

LOCATION: (4827)...(4925)

NAME/KEY: CDS

LOCATION: (6322)...(6447)

NAME/KEY: CDS

LOCATION: (6543)...(6638)

NAME/KEY: CDS

LOCATION: (7137)...(7316)

NAME/KEY: CDS

LOCATION: (7439)...(7553)

US-09-167-681-45

Query Match

Best Local Similarity 2.5%; Score 51; DB 4; Length 8453;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1169 ttttttttgagacagagctcactctgttgcacagctggagtgcaagtgg 1219
|||||
Db 1709 ttttttttgagacagagctcactctgttgcacagctggagtgcaagtgg 1759

RESULT 11

US-09-009-913-1

Sequence 1, Application US/09009913

Patent No. 6087485

GENERAL INFORMATION:

APPLICANT: Axys Pharmaceuticals, Inc.

TITLE OF INVENTION: Asthma Related Genes

NUMBER OF SEQUENCES: 339

CORRESPONDENCE ADDRESS:

ADDRESSEE: Bozicevic & Reed, LLP

STREET: 285 Hamilton Ave, Suite 200
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94301

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/009,913

FILING DATE: 21-JAN-1998

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Sherwood, Pamela J

REGISTRATION NUMBER: 36,677

REFERENCE/DOCKET NUMBER: SEQ-4P

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-327-3231

TELEFAX: 650-327-3231

TELEX:

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 72928 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: Genomic DNA

US-09-009-913-1

Query Match 2.3%; Score 46; DB 3; Length 72928;
Best Local Similarity 100.0%; Pred. No. 7.4e-11;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1180 gacagagctcactctgttgcacagctggagtgcaagtggcacaat 1225
|||||

Db 9587 CACAGAGTCTCAGTCTGTGTCACAGGCTGGAGTGCAGTGCACAAT 9632

RESULT 12

US-08-611-587-1/c

Sequence 1, Application US/08611587

Patent No. 6150091

GENERAL INFORMATION:

APPLICANT: PANDOLFO, MASSIMO

APPLICANT: MONTERMINI, LAURA

APPLICANT: MOLTO, MARIA D.

APPLICANT: Koenig, Michael

APPLICANT: Campuzano, Victoria

APPLICANT: Cossee, Mireille

TITLE OF INVENTION: Direct Diagnosis of Friedreich Ataxia

NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fulbright & Jaworski L.L.P. Patent Dept.

STREET: 1301 McKinney, Suite 5100

CITY: Houston

STATE: Texas

COUNTRY: U.S.

ZIP: 77010

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/611,587

FILING DATE: 03-MAR-1996

CLASSIFICATION: 436

ATTORNEY/AGENT INFORMATION:
NAME: Brashears-Macatee, Sarah J.
REGISTRATION NUMBER: 38,087
REFERENCE/DOCKET NUMBER: D-5901
TELEPHONE: 713-651-5620
TELEFAX: 713-651-5246
TELEX: 76-2829

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 8353 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapien
POSITION IN GENOME:
CHROMOSOME/SEGMENT: 9q13
UNITS: bp
US-08-611-587-1

Query Match 2.2%; Score 45; DB 3; Length 8353;
Best Local Similarity 100.0%; Pred. No. 2.3e-10;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1175 tttagacagagctcactctgttgcaggctgagtgagtgagtg 1219
|||||

Db 2090 TTTGACACAGCTCACTCTGTGTTGCCAGGCTGGAGTGCAGTGG 2046

RESULT 13

US-08-370-319C-12/c
Sequence 12, Application US/08370319C
Patent No. 5856091

GENERAL INFORMATION:
APPLICANT: Brichard, Vincent; Van Pel, Aline;
APPLICANT: Traversari, Catia; W lfel, Thomas; Coullie, Pierre;
APPLICANT: Boon-Falleur, Thierry; De Plaeen, Etienne.
TITLE OF INVENTION: ISOLATED NUCLEIC ACID SEQUENCE CODING FOR A
TITLE OF INVENTION: TUMOR REJECTION ANTIGEN PERCURSOR PROCESSED TO AT LEAST ONE TU
TITLE OF INVENTION: REJECTION ANTIGEN PRESENTED BY HLA-A2

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Felfe & Lynch

STREET: 805 Third Avenue

CITY: New York City

STATE: New York

COUNTRY: USA

ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage

COMPUTER: IBM PS/2

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/370,319C

FILING DATE: 10-JANUARY-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/272,351

FILING DATE: 8-JULY-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/032,978

FILING DATE: 18-MAR-1993

ATTORNEY/AGENT INFORMATION:

NAME: Hanson, No. 5856091man D.

REGISTRATION NUMBER: 30,946

REFERENCE/DOCKET NUMBER: LUD 5377.1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 4129 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

FEATURE:

OTHER INFORMATION: The sequence is preceded by an

SEQUENCE CHARACTERISTICS: unsequenced portion of from 4.7 to 5.3

OTHER INFORMATION: kilobases

US-08-370-319C-12

Query Match 2.2%; Score 44; DB 2; Length 4129;

Best Local Similarity 100.0%; Pred. No. 6.3e-10;

Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1169 ttttttttgacagagctcactctgttgcaggctgagtg 1212
|||||

Db 1516 TTTTGTGACAGAGCTCACTCTGTGTTGCCAGGCTGGAGT 1473

RESULT 14

US-09-224-834-12/c

Sequence 12, Application US/09224834

Patent No. 6201111

GENERAL INFORMATION:

APPLICANT: Brichard, Vincent; Van Pel, Aline;

APPLICANT: Traversari, Catia; W lfel, Thomas; Coullie, Pierre;

APPLICANT: Boon-Falleur, Thierry; De Plaeen, Etienne

TITLE OF INVENTION: ISOLATED NUCLEIC ACID SEQUENCE CODING FOR A

TITLE OF INVENTION: TUMOR REJECTION ANTIGEN PERCURSOR PROCESSED TO AT LEAST ONE

TITLE OF INVENTION: REJECTION ANTIGEN PRESENTED BY HLA-A2

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Felfe & Lynch

STREET: 805 Third Avenue

CITY: New York City

STATE: New York

COUNTRY: USA

ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage

COMPUTER: IBM PS/2

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/224,834

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/370,319

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/032,978

FILING DATE: 18-MAR-1993

ATTORNEY/AGENT INFORMATION:

NAME: Hanson, No. 6201111man D.

REGISTRATION NUMBER: 30,946

REFERENCE/DOCKET NUMBER: LUD 5377.1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 688-9200

TELEFAX: (212) 838-3884

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 4129 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

FEATURE:

OTHER INFORMATION: The sequence is preceded by an

OTHER INFORMATION: unsequenced portion of from 4.7 to 5.3
OTHER INFORMATION: kilobases
US-09-224-834-12

Query Match 2.2%; Score 44; DB 4; Length 4129;
Best Local Similarity 100.0%; Pred. No. 6.3e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1169 ttttttttgagacagagctcactctgtgtccaggctgaggt 1212
|||||
Db 1516 TTTTGTGAGACAGAGTCTCACTCTGTGCCAGGCTGGAGT 1473

RESULT 15

US-09-128-155-17/c
; Sequence 17, Application US/09128155
; Patent No. 6117654
; GENERAL INFORMATION:
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
; FILE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: 09404/052001
; CURRENT APPLICATION NUMBER: US/09/128,155
; CURRENT FILING DATE: 1998-08-03
; EARLIER APPLICATION NUMBER: US 60/091,650
; EARLIER FILING DATE: 1998-07-02
; EARLIER APPLICATION NUMBER: US 60/054,646
; EARLIER FILING DATE: 1997-08-04
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 176373
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(176373)
; OTHER INFORMATION: n = A,T,C or G
US-09-128-155-17

Query Match 2.2%; Score 44; DB 3; Length 176373;
Best Local Similarity 100.0%; Pred. No. 5.1e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1184 gagtctcactgtgtgccaggctggagtcagtgaggcacaatct 1227
|||||
Db 172862 GAGTCTCACTGTGTGCCAGGCTGGAGTCAGTGGCACAATCT 172819

Search completed: October 7, 2002, 12:56:52
Job time: 6736 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 7, 2002, 11:08:30 ; Search time 238.5 Seconds
(without alignments)
14519.985 Million cell updates/sec

Title: US-09-813-492-1

Perfect score: 2017

Sequence: 1 tagataccctggaacacctcc.....ataccataaaaaaaaaa 2017

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 1736436 seqs, 858457221 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_032802.*

1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	498	24.7	731 19	AAV38291 Human C-C chemokine
2	498	24.7	731 21	AAA47545 Primate Vic nucleos
3	498	24.7	778 22	AAF59231 Human G-protein co
4	434	21.5	643 22	AAH34102 Human colon cancer
5	403	20.0	698 22	AAH99420 Human protein enco
6	382	18.9	383 20	AAH40376 Human secreted pro
7	382	18.9	566 21	AAA52456 cDNA encoding huma
8	336	16.7	675 23	AAH76558 DNA encoding novel
9	313	15.5	3117 20	AAH89389 Human mammary asso

10	279	13.8	381	20	AAH89390	Human mammary asso
11	221	11.0	294	22	AAH89722	Human Rintin CD
12	202	10.0	311	20	AAH89394	Human MACK hybrid
13	192	9.5	474	23	AAH81262	DNA encoding novel
14	180	8.9	253	21	AAH91767	Human breast speci
15	126	6.2	254	20	AAH51656	Human secreted pro
16	104	5.2	104	20	AAH89391	Human mammary asso
17	91	4.5	445	19	AAH38293	Human C-C chemokine
18	69	3.4	496	19	AAH38292	Human C-C chemokine
19	64	3.2	486	23	AAH81261	DNA encoding novel
20	60	3.0	173	22	AAH75733	Human immune/haema
21	57	2.8	481	22	AAH18403	Human nervous syst
22	54	2.7	2264	22	AAH17045	Human cDNA sequenc
23	54	2.7	2854	22	AAH05517	Human reproductive
24	54	2.7	3271	22	AAH05518	Human reproductive
25	54	2.7	10122	22	AAH87363	Human immune/haema
26	54	2.7	11484	22	AAH87360	Human immune/haema
27	54	2.7	11497	22	AAH87362	Human immune/haema
28	54	2.7	15536	22	AAH87354	Human immune/haema
29	54	2.7	20645	22	AAH05355	Human reproductive
30	54	2.7	23241	22	AAH97870	Human neuroblastom
31	54	2.7	23241	22	AAH97871	Human neuroblastom
32	53	2.6	1467	22	AAH86144	Human immune/haema
33	53	2.6	15914	22	AAH84889	Human immune/haema
34	53	2.6	23823	22	AAH79160	Human immune/haema
35	53	2.6	23825	22	AAH79161	Human immune/haema
36	53	2.6	31304	22	AAH08089	Human ovarian and
37	53	2.6	31304	22	AAH08088	Human reproductive
38	52	2.6	88	22	AAH77683	Human immune/haema
39	52	2.6	101	22	AAH30479	DNA encoding novel
40	52	2.6	101	22	AAH06259	Human reproductive
41	52	2.6	108	22	AAH02850	Human reproductive
42	52	2.6	130	22	AAH77684	Human immune/haema
43	52	2.6	191	21	AAH21190	Human secreted pro
44	52	2.6	300	20	AAH98284	Human cancer cell
45	52	2.6	333	22	AAH90323	Human polynucleoti

ALIGNMENTS

RESULT 1

AAV38291

ID AAV38291 standard; cDNA; 731 BP.

XX AAV38291;

XX 12-OCT-1998 (first entry)

DT Human C-C chemokine DVic-1 cDNA.

DE DVic-1; DNAX Vic-1; C-C chemokine; cytokine; human; immune system;

KW cancer; cell proliferation; therapy; diagnosis; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 56..439

FT sig_peptide /*tag= a

FT /*tag= b

FT mat_peptide 122..436

FT /*tag= c

XX WO9823750-A2.

PN 04-JUN-1998.

XX 26-NOV-1997;

XX 97WO-US21092.

XX 05-DEC-1996;

XX 96US-0761071.

XX 27-NOV-1996;

XX 96US-0031805.

Db 122 agccatactcccattgctccagctgtgtgacgaggtttccatcatcatattccagaag 181
Qy 268 qctcctggaagagtgaatatgtgtcgcacatccagagagctgagtgaggattgtgacttgc 327
Db 182 gctcctggaagagtgaatatgtgtcgcacatccagagagctgagtgaggattgtgacttgc 241
Qy 328 tgcgtcatccttcacgtcaagcgacgaagaatctgtgtcagccgcacacaccatactgt 387
Db 242 tgcgtcatccttcacgtcaagcgacgaagaatctgtgtcagccgcacacaccatactgt 301
Qy 388 taagcagtgatgaagtgcaagtgccagaaataatggttaaaggaaatgtttgcccacag 447
Db 302 taagcagtgatgaagtgcaagtgccagaaataatggttaaaggaaatgtttgcccacag 361
Qy 448 gaagaaacaccatggcaagagg 469
Db 362 gaagaaacaccatggcaagagg 383

RESULT 7
AAA52456
ID AAA52456 standard; cDNA; 566 BP.
XX AC AAA52456;
XX DT 25-SEP-2000 (first entry)
XX DE cDNA encoding human growth factor related molecule GFRP-2.
XX KW Human GFRP-2; growth factor related molecule; breast tissue;
KW breast tumour; CC chemokine; hTECK homologue; developmental disorder;
KW cell proliferative disorder; immune disorder; reproductive disorder;
KW cardiovascular disorder; bacterial infection; viral; fungal; parasitic;
KW cancer; allergy; asthma; arteriosclerosis; therapy; diagnosis;
KW consensus; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT CDS 50..433
FT /*tag= a
FT /product= "Human GFRP-2"
XX PN WO200024774-A2.
XX PD 04-MAY-2000.
XX PF 28-OCT-1999; 99NO-US25458.
XX PR 28-OCT-1998; 98US-0181711.
PR 11-DEC-1998; 98US-0209547.
PR 17-MAY-1999; 99US-0313457.
XX PA (INCY-) INCYTE PHARM INC.
XX PI Tang YT, Yue H, Hillman JL, Corley NC, Guegler KJ, Baughn MR;
PI Au-Young J;
XX WP1: 2000-350695/30.
DR P-PSDB; AAB03001.
XX
XX Human growth factor related molecule protein useful for the diagnosis
PT and treatment of disorders associated with its activity including
PT developmental, cell proliferative, immune, reproductive and
PT cardiovascular disorders and infections -
XX Claim 9; Page 74; 80pp; English.
XX PS
XX This sequence represents cDNA encoding human growth factor related
CC molecule GFRP-2. cDNA encoding GFRP-1 was initially identified in a
CC breast tissue cDNA library, and the present sequence represents a
CC consensus derived from several overlapping and/or extended cDNA clones.
CC GFRP-2 is probably a CC chemokine and has chemical and structural

CC homology with hTECK (20% identity). GFRP-2 and hTECK are both basic
CC proteins, having isoelectric points of 10.1 and 10.2, respectively.
CC GFRP-2 was found by Northern analysis to be expressed in both tumorous
CC and nontumorous breast tissue. GFRP proteins (AAB03000-B03003),
CC nucleotides encoding them (AAA52455-A52458), GFRP agonists and
CC antagonists may be used to treat a wide variety of diseases associated
CC with increased or decreased expression or activity of GFRP proteins.
CC Conditions which may be treated include developmental disorders, cell
CC proliferative disorders (e.g., cancers). Immune disorders (e.g.,
CC allergies, asthma), reproductive disorders (e.g., menstrual cycle
CC disorders) cardiovascular disorders (e.g., arteriosclerosis) and
CC bacterial, viral, fungal or parasitic infections. Additionally, GFRP
CC proteins and nucleotides can be used in the diagnosis of such disorders.
XX
SQ Sequence 566 BP; 152 A; 137 C; 144 G; 133 T; 0 Other;
Query Match 18.9%; Score 382; DB 21; Length 566;
Best Local Similarity 99.8%; Pred. No. 2.8e-146;
Matches 432; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 97 tcgaacagcctcacctgtgttctgtcagtcagtagggcaggcagggaatcgacgagag 156
Db 1 tcgaacagcctcacctgtgttctgtcagtcagtagggcaggcagggaatcgacgagag 60
Qy 157 aggaactcgccatcgctggcctgtgtctgtcgcgcctcacatgctcagaagccatact 216
Db 61 aggaactcgccatcgctggcctgtgtctgtcgcgcctcacatgctcagaagccatact 120
Qy 217 tccattgctcctcagctgtgtgcacgaggtttcacatcatatttccagaagctcctgga 276
Db 121 tccattgctcctcagctgtgtgcacgaggtttcacatcatatttccagaagctcctgga 180
Qy 277 aagagtgaatatgtgtcgcacatcagagaactatgggagttgtgacttgctcgtctcat 336
Db 181 aagagtgaatatgtgtcgcacatcagagaactatgggagttgtgacttgctcgtctcat 240
Qy 337 ccttcattgcaagcgacagaagaatctgtcagccgcacacacatactgttaagcagtg 396
Db 241 ccttcattgcaagcgacagaagaatctgtcagccgcacacacatactgttaagcagtg 300
Qy 397 gatgaagtgcagctgccaaagaaaatggttaaaggaaatgtttgccacagagaagaaaca 456
Db 301 gatgaagtgcagctgccaaagaaaatggttaaaggaaatgtttgccacagagaagaaaca 360
Qy 457 ccattggcaagaggaacagtaacagggcacatcaggggaaacacacatacggccataa 516
Db 361 ccattggcaagagggacagtaacagggcacatcaggggaaacacacatacggccataa 420
Qy 517 aactccttattag 529
Db 421 aactccttattag 433

RESULT 8
AAS76558/c
ID AAS76558 standard; cDNA; 675 BP.
XX AC AAS76558;
XX DT 13-FEB-2002 (first entry)
XX DE DNA encoding novel human diagnostic protein #12362.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX


```
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG12371.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID No 12362; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 675 BP; 152 A; 166 C; 150 G; 207 T; 0 other;

Query Match 16.7%; Score 336; DB 23; Length 675;
Best Local Similarity 100.0%; Pred. No. 1.6e-127; Indels 0; Gaps 0;
Matches 336; Conservative 0; Mismatches 0;

QY 366 tcagccgcacacaccatactgttaagcagtgatgaagtgcaagtcgcaagaaaaatg 425
DB 364 TCAGCCCGCACACCACTACTGTTAGCAGTGGATGAAGTGGCAAGCTGCCAAGAAAATG 305
QY 426 gtaaggaaatgtttgccacaggaagaacacccatggcaaggaggaacagtaacagggcac 485
DB 304 GTAAGGAAATGTTTGGCCACAGGAAGAACACCATGGCAAGAGGAAACAGTAACAGGGCAC 245
QY 486 atcaggggaacacagaacatcagcgcataaaactccttatttagagagctctacagataaa 545
DB 244 ATTCAGGGGAACACAGAAACATACGCCCAATAAAACCTTATTAGAGAGTCTACAGATAAA 185
QY 546 tctcagagacaaatcctcaagtgcactggccatgattggttgtaagtttatctctga 605
DB 184 TCTACAGAGACAAATTCCTCAAGTGAGCTGGCCATGTTGTTGAAGTTTATCATCTGA 125
QY 606 attctccttattgtagacaagacaacaaataattggtttttaaaaaatgaacaaat 665
DB 124 ATTCCTCTTATTGTAGACAACAGAACAAACAAATATTGGTTTAAAAAATGAACAA 65
QY 666 tctgcgggtatgcaaatgtagccaataataactactaa 701
DB 64 TGTGCGGTATGCAAAATGTAGCCAATTAATATACTCAA 29
```

```
RESULT 9
AAX89389
ID AAX89389 standard; cDNA; 3117 BP.
XX
AC AAX89389;
XX
DT 29-SEP-1999 (first entry)
XX
DE Human mammary associated chemokine (MACK) encoding cDNA.
XX
KW Chemokine; breast tissue; breast milk; breast disease; vaccine; human;
KW inflammation; infection; mastitis; benign cystitis; hyperplasia;
KW mammary associated chemokine; MACK; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 47..430
FT /tag= a
FT /transl_except= (pos: 254..256, aa: Xaa)
FT /transl_except= (pos: 317..319, aa: Xaa)
FT /product= "MACK"
FT /note= "Xaa = unknown"
FT sig_peptide 47..115
FT mat_peptide 116..430
FT /tag= b
FT /tag= c
XX
XX WO9936540-A1.
XX
XX 22-JUL-1999.
XX
XX 12-JAN-1999; 99WO-US00651.
XX
XX 09-JUL-1998; 98US-0092155.
XX
XX 20-JAN-1998; 98US-0071899.
XX
XX (CODO-) CODON DIAGNOSTICS LLC.
XX
XX Dyster LM, Frustaci JM, Papsidero LD;
XX
XX WPI; 1999-458469/38.
XX
XX P-PSDB; AAY29092, AAY29093.
XX
XX A mammary associated chemokine and related polynucleotides, useful
XX for detection and treatment of breast disease, especially cancer
XX
XX Claim 19; Page 18-20; 76pp; English.
XX
XX The invention provides an isolated human chemokine, which is
XX preferentially expressed in breast tissue or detected in breast milk. An
XX antibody that recognizes the novel chemokine, or a chemokine-derived
XX antigenic peptide, can be used to treat breast disease in a patient. A
XX peptide, which binds to a cellular receptor for the chemokine, can also
XX be used to treat breast disease. Antigenic peptides of the chemokine can
XX be used to vaccinate patients against breast disease. The chemokine
XX polynucleotide sequences and the chemokine protein can be detected in
XX samples with primers, probes and antibodies using standard techniques.
XX This is useful for detecting breast disease. Other breast diseases that
XX may be treated or detected with the chemokine and its encoding
XX polynucleotides include inflammations, infections, mastitis, benign
XX cystitis, and benign hyperplasias as well as other malignancies. The
XX present sequence represents the nucleotide sequence of the human mammary
XX associated chemokine (MACK).
XX
XX Sequence 3117 BP; 860 A; 582 C; 566 G; 921 T; 188 other;

Query Match 15.5%; Score 313; DB 20; Length 3117;
Best Local Similarity 100.0%; Pred. No. 3.4e-118;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 697 ctcaaaactcctgggctcaagcgatcctccacaccttagctcccaagactgggattata 756
```


THIS PAGE BLANK (uspto)

ORIGIN

Query Match 28.1%; Score 567; DB 10; Length 698;
Best Local Similarity 99.8%; Pred. No. 5.le-160;
Matches 617; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1290 actgattacagggccagccagccacacaccccgctgattttgtatttttagtagagac 1349
|||||
Db 1 ACTGATATACAGGGCCCGCCAGCCACACACCCGCTGATTTTGTATTTTAGTAGAC 60
Qy 1350 ggggttttccacgttgccggggtggtctcaaaactcttgacctcaagtgaaccacccg 1409
|||||
Db 61 GGGGTTTTCCACGTTGGCGGGTGGTCTCAAACTCTTGACCTCAAGTGAACCCCGC 120
Qy 1410 ctgtacctcccaaaactgtaattaccagctgaagccacacatccgggctcacagctt 1469
|||||
Db 121 CTTGTCCTCCCAAGAGTCTGGAAATACCCAGGCTGAGCCACCATGCCGGGCTCACACGTTT 180
Qy 1470 ggttgataccattgtgccattcctcttttggcctctttttgtccatagaggttcaag 1529
|||||
Db 181 GAGTTGATACCATTTGCGCATCTCTTTTGGCCTCTTTTGTCCATAGAGGCTTCAAG 240
Qy 1530 atagataggttaagaccagtagtgttccataagaagccaatagagagcaggccacttt 1589
|||||
Db 241 ATAGATAGGTAAGAGCCCGAGTAGTGTTCATAAGAAGCCCAATAGAGAGCAGGCGCACTTT 300
Qy 1590 atcaagtgagcaggtgtccggggtccctgctgctagtcacaaagcgtgtgttgcag 1649
|||||
Db 301 ATCAGGTGGCAGGTGCTCGGGGCTCCCTGCTGGCTAGTCCCAAGCGGTGTGTGGCAG 360
Qy 1650 gatgtctggaggtgataatgggacacagagcagcagctgctccataggttataaagtc 1709
|||||
Db 361 GATGCTTTGGAGTGATATATGGGACACACAGAGGCACCTGAGTCTCCATAGGTTAAATGC 420
Qy 1710 caccaaaactggccttgcctaataatccctcattgactatttagcattattattttat 1769
|||||
Db 421 CACCAAAAGTGGCTTTGCCTAAATCCCTCATTTAGCTATTATTTATTTATTTAT 480
Qy 1770 ttctctgacattctgcaagctttgtattattatttcaacttattagtagtgaggaatttg 1829
|||||
Db 481 TTTCTCTGACATTTCTGCAAGCTTTGTATTTATTTATTTCTTATTTATTTATTTATTT 540
Qy 1830 aggcctcttagaggttaaaatgacttgcacaggttcacacaggaagtgtgcagagacaaagcttt 1889
|||||
Db 541 AGGCTCTTAGAGTAAATGACTTCCCGAGGTCACACAGGAGTGGCAGAGACACAGCTTT 600
Qy 1890 ttaataagaaaaatta 1907
|||||
Db 601 TTAATAAAGAAAAATTA 618

RESULT 2
AW469978/c
LOCUS xr27a11.x1 NCI_CGAP_Ut4 Homo sapiens cDNA clone IMAGE:2761340 3',
DEFINITION mRNA sequence.
ACCESSION AW469978
VERSION AW469978.1 GI:7040084
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 508)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution Information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Possible reversed clone: polyt not found
Seq primer: -40UP from Gibco
High quality sequence stop: 408.

FEATURES
source

1..508
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2761340"
/clone_lib="NCI_CGAP_Ut4"
/tissue_type="serous papillary carcinoma, high grade, 2
pooled tumors"
/lab_host="DH10B"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.48 kb. Life Technologies catalog #:
11542-016"

BASE COUNT 150 a 102 c 92 g 164 t
ORIGIN

Query Match 25.2%; Score 508; DB 9; Length 508;
Best Local Similarity 100.0%; Pred. No. 2.8e-142;

Matches 508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1495 ctttgtgcctctttttgtccatagagcgttccaagatagatagtagtaagagcccgtagtg 1554
|||||

Db 508 CTTTGTGGCCTCTTTTGTCCATAGAGGCTTCAAGATAGATAGTAGTAGAGCCCGAGTAGTG 449
|||||

Qy 1555 ttcaataagaagccaatagagcaggagccactttatcaggttgcccggttcccggtcct 1614
|||||

Db 448 TTCATAAGAAGCCAAATAGAGAGCAGGAGCCACTTTATCAGGTGCGAGGTGTCCCGGCGCT 389
|||||

Qy 1615 ccctgtcgttagtcccaagcgggtgttgcagagatgtcttgaggtgataatgggac 1674
|||||

Db 388 CCCGTCTGGCTAGTCCCAAGCGGTGGTGTGCCAGGATGTCTTGGAGGTGATATATGGGAC 329
|||||

Qy 1675 acacagaggcactgagtctccataggttaaaatgccaccacaaactggcctttgcctaata 1734
|||||

Db 328 ACACAGAGGCACCTCAGTCTCCATAGGTTAAATGCCACCAAAAGTGGCCTTTGCCTTAATA 269
|||||

Qy 1735 tccctcattgactatttagcatttaattttatttttccctgacatttctgcagcgttg 1794
|||||

Db 268 TCCTCATGACTATTAGCATTTAATTTATTTATTTTCTCTGACATTTCTGCAAGCTTTG 209
|||||

Qy 1795 tatttatatttccactttatagatgaggaatttgagcctcttagaggttaaaatgacttg 1854
|||||

Db 208 TATTTATTTTCCACTTTATAGATGAGGAATTTGAGGCTCTTAGAGGTAAATGACTTG 149
|||||

Qy 1855 cccaggtcacacaggaagtggcagagacaagctttttaaataagaaaaaattataaataa 1914
|||||

Db 148 CCCAGGTCACACAGAGAGTGGCAGAGACAGCTTTTAAATAAGAAAAAATTAATAAAT 89
|||||

Qy 1915 ataataagagtaacttaaaatattataaacacacaaattttaaatttaaacccgtgat 1974
|||||

Db 88 ATAATATGAGAGTAACCTAAATAATTAAATAAACCCACAAATTTAAATTAATTAACCGTAT 29
|||||

Qy 1975 aaccacaatttaataaagttaagatacc 2002
|||||

Db 28 AACCAACATTAAATAAAGTTAAGATACC 1
|||||

RESULT 3

BC530240
LOCUS BC530240 869 bp mRNA linear EST 03-APR-2001
DEFINITION 602559348F1 NIH_MGC_61 Homo sapiens cDNA clone IMAGE:4697338 5',
mRNA sequence.


```
ACCESSION      BG530240
KEYWORDS       BG530240.1  GI:13521777
VERSIONS       EST.
SOURCE         human.
ORGANISM       Homo sapiens

REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS        Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE          1 (bases 1 to 869)
JOURNAL        NIH-MGC http://mgc.nci.nih.gov/
COMMENT        National Institutes of Health, Mammalian Gene Collection (MGC)
               Unpublished (1999)
               Contact: Robert Strausberg, Ph.D.
               Email: cgapbs-r@mail.nih.gov
               Tissue Procurement: ATCC
               CDNA Library Preparation: CLONETECH Laboratories, Inc.
               CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
               DNA Sequencing by: Incyte Genomics, Inc.
               Clone distribution: MGC clone distribution information can be
               found through the I.M.A.G.E. Consortium/LLNL at:
               http://image.llnl.gov
               Plate: L1CM1525 row: p column: 11
               High quality sequence stop: 748.
               Location/Qualifiers
               1..869
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:4697338"
                /clone_lib="NIH_MGC_61"
                /tissue_type="embryonal carcinoma"
                /lab_host="DH10B (T1 phage-resistant)"
                /note="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1:
                SfiI (ggcgcttcggcc); Site_2: SfiI (ggccatttgccc);
                Double-stranded cDNA was prepared from cell line RNA. 5'
                and 3' adaptors were used in cloning as follows: 5'
                adaptor sequence: 5'-ATTCTAGAGCGGCGGCACATG-dn(30)BN-3'
                (where B = A, C, G and N = A, C, G, or T). Average
                insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies
                contained inserts by PCR. This library was enriched for
                full-length clones and was constructed by Clontech
                Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
                Library."
BASE COUNT     261 a 185 c 215 g 208 t
ORIGIN

Query Match    24.5%; Score 494; DB 10; Length 869;
Best Local Similarity 100.0%; Pred. No. 2.7e-138;
Matches 494; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 208 agccatacttccattgcttcagctgttgacgaggtttcacatcatatttcagaag 267
      |||||||
Db 94  agccatacttccattgcttcagctgttgacgaggtttcacatcatatttcagaag 153

Qy 268 gctctggaaagagtgaatatgttcgcattccagagctgattgggattgtgactggc 327
      |||||||
Db 154 gctctggaaagagtgaatatgttcgcattccagagctgattgggattgtgactggc 213

Qy 328 tgcgtcatcttcatgtccaagcagagaatctgtgtcgccgcgaaccatactcgt 387
      |||||||
Db 214 tgcgtcatcttcatgtccaagcagagaatctgtgtcgccgcgaaccatactcgt 273

Qy 388 taagcagtgatgaagtgcaagctgcccaagaaaatgtaaggaaatgttgccacag 447
      |||||||
Db 274 taagcagtgatgaagtgcaagctgcccaagaaaatgtaaggaaatgttgccacag 333

Qy 448 gaagaaacacacattggcaagagaacagtaacaggggcacatcaggggaaacacgaaacata 507
      |||||||
Db 334 gaagaaacacacattggcaagagaacagtaacaggggcacatcaggggaaacacgaaacata 393

Qy 508 cggccataaaactcttattagagagtctacagataaattctacagagacaattctcaag 567
      |||||||
Db 394 cggccataaaactcttattagagagtctacagataaattctacagagacaattctcaag 453
```

```
Qy 568 tggacttgccatgattgttgtaagtttatacatcattctcctattattgtagaaca 627
      |||||||
Db 454 tggacttgccatgattgttgtaagtttatacatcattctcctattattgtagaaca 513

Qy 628 gaacaaacaaataattggttttttaaaaaaagaacaattgtcggtatgcaaatagcc 687
      |||||||
Db 514 gaacaaacaaataattggttttttaaaaaaagaacaattgtcggtatgcaaatagcc 573

Qy 688 ataataatactcaa 701
      |||||||
Db 574 ataataatactcaa 587

RESULT 4
BG682175      BG682175      816 bp      mRNA      linear      EST 01-MAY-2001
LOCUS         602629474f1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4754250 5',
DEFINITION    mRNA sequence.
ACCESSION     BG682175
VERSION       BG682175.1  GI:13913572
KEYWORDS      EST.
SOURCE        human.
ORGANISM      Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1 (bases 1 to 816)
AUTHORS       NIH-MGC http://mgc.nci.nih.gov/
TITLE         National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL       Unpublished (1999)
COMMENT       Contact: Robert Strausberg, Ph.D.
               Email: cgapbs-r@mail.nih.gov
               Tissue Procurement: James Cleaver, M.D.
               CDNA Library Preparation: Life Technologies, Inc.
               CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
               DNA Sequencing by: Incyte Genomics, Inc.
               Clone distribution: MGC clone distribution information can be
               found through the I.M.A.G.E. Consortium/LLNL at:
               http://image.llnl.gov
               Plate: L1AM10615 row: c column: 19
               High quality sequence stop: 766.
               Location/Qualifiers
               1..816
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:4754250"
                /clone_lib="NCI_CGAP_Skn4"
                /tissue_type="squamous cell carcinoma"
                /lab_host="DH10B (T1 phage-resistant)"
                /note="Organ: skin; Vector: pCMV-SpORF6; Site_1: NotI;
                Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                Average insert size 1.5kb. Library constructed by Life
                Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT    184 a 199 c 172 g 261 t
ORIGIN

Query Match    20.8%; Score 420; DB 10; Length 816;
Best Local Similarity 99.7%; Pred. No. 3.6e-116;
Matches 660; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

Qy 802 ttcaggtttaaattgttttgtaagaatttctacgtgaattcttgactattttgtca 861
      |||||||
Db 32  ttcaggtttaaattgttttgtaagaatttctacgtgaattcttgactattttgtca 91

Qy 862 tttagagttcataaattattaggggtttattttctaaataagaattttaaacataataaa 921
      |||||||
Db 92  tttagagttcataaattattaggggtttattttctaaataagaattttaaacataataaa 151

Qy 922 cttcaaaacgtctagtttgtagtagctaccgtgtgttttgattgaaatttctgatactgaa 981
      |||||||
Db 152 cttcaaaacgtctagtttgtagtagctaccgtgtgttttgattgaaatttctgatactgaa 211
```



```

RESULT 8
BE463561/c
LOCUS
DEFINITION
hw25b05.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:3183921 3',
mRNA sequence.
ACCESSION
BE463561
VERSION
BE463561.1 GI:9509336
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 362)
AUTHORS
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmett-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
Seq primer: -400P from Gibco.
Location/Qualifiers
1. 362
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3183921"
/clone_lib="NCI_CGAP_Kid11"
/lab_host="DH10B"
/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
Plasmid DNA from the normalized library NCI_CGAP_Kid3 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneIDs 132376-1323911, 1456007-1456775, and
1500552-1502855). Subtraction by Bento Soares and M.
Fatima Bonaldo."
BASE COUNT 107 a 81 c 67 g 107 t
ORIGIN
Query Match 16.1%; Score 324; DB 10; Length 362;
Best Local Similarity 100.0%; Pred. No. 3.5e-87;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1609 gggcctccctgctgtagtcccaagcgggtgtgtgccaggatgtcttgagggtgataa 1668
|||||
Db 324 GGGCCTCCCTGCTGCTAGTCCCAAGCGGTGGTGTGCCAGGATGCTTGGAGGTGATAA 265
|||||

Qy 1669 tgggacacacagagcactgagttccatcagggttaaatgccaccacaaactggcctttgc 1728
|||||
Db 264 TGGGACACACAGAGGCACTGAGTCTCCATAGGTTAAATGCCACCAAACTGGCCTTTGC 205
|||||

Qy 1729 ctaatatccctcattgactatttagcatttaattttatttctcgtgacattctgcaa 1788
|||||
Db 204 CTAAATATCCCTCATTTGACTATTATAGCATTTAAATTATTTATTTCTCGACATTTCTGCAA 145
|||||

Qy 1789 gctttgtatttatatttccactttatagatgaggaaatttgaggctctcttagaggtaaaaa 1848
|||||
Db 144 GCTTTGTATTATATTTCCACTTTATAGATGAGAAATTTTGGGCTCTTAGAGGTAAAT 85
|||||

Qy 1849 gacttggccagggtcacacaggaagtggcagagacaagcttttttaataagaacaaataa 1908
|||||

```

```

Db 84 GACTTGCCCGAGTCCACAGAGAGTGGCAGAGACGCTTTTAAATAAGAAAAATTAA 25

Qy 1909 taaaataataatagagagtaactt 1932
|||||
Db 24 TAAATATAATATGAGAGTAACCT 1

RESULT 9
N63913/c
LOCUS
DEFINITION
za27h10.sl Soares fetal liver spleen lNFLS Homo sapiens cDNA clone
IMAGE:293827 3' similar to contains element MIR repetitive element
; mRNA sequence.
ACCESSION
N63913
VERSION
N63913.1 GI:1211742
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 577)
AUTHORS
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman
, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,
Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston
, R., Williamson, A., Wohldmann, P. and Wilson, R.
TITLE
The WashU-Merck EST Project
JOURNAL
Unpublished (1995)
COMMENT
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: ml3 -40 forward
High quality sequence stop: 356.
Location/Qualifiers
1. 577
/organism="Homo sapiens"
/db_xref="GDB:3801674"
/db_xref="taxon:9606"
/clone="IMAGE:293827"
/clone_lib="Soares fetal liver spleen lNFLS"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
[5' AACTGGAAGATTAAATTAAGATCTTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 168 a 135 c 132 g 141 t
ORIGIN
Query Match 14.3%; Score 289; DB 10; Length 577;
Best Local Similarity 100.0%; Pred. No. 6.6e-77;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1609 gggcctccctgctgtagtcccaagcgggtgtgtgccaggatgtcttgagggtgataa 1668
|||||
Db 289 GGGCCTCCCTGCTGCTAGTCCCAAGCGGTGGTGTGCCAGGATGCTTGGAGGTGATAA 230
|||||

Qy 1669 tgggacacacagagcactgagttccatcagggttaaatgccaccacaaactggcctttgc 1728
|||||
Db 229 TGGGACACACAGAGGCACTGAGTCTCCATAGGTTAAATGCCACCAAACTGGCCTTTGC 170
|||||

Qy 1729 ctaatatccctcattgactatttagcatttaattttatttctcgtgacattctgcaa 1788
|||||

```

```

|||||
Db 169 CTAATACCTCATTTGACTATTTAGCATTTAAATTTATTTTCTCGACATTTTCGCA 110
|||||
Qy 1789 gctttgattatttccactttatagatgaggaatttagagctcttagaggtaaaaat 1848
|||||
Db 109 GCTTTGATTTATTTCCACTTTATAGATGAGGAATTTGAGGCTCTTAGAGGTAAAT 50
|||||
Qy 1849 gacttccagggtcacacagggaagtgcagagacaagcttttaataa 1897
|||||
Db 49 GACTTCCAGGTCACACAGGAAGTGCAGACAGCTTTTAAATAA 1
|||||

RESULT 10
N98285
LOCUS
DEFINITION
N98285 437 bp mRNA linear EST 10-APR-1996
za27h10.r1 Soares fetal liver spleen lNFLS Homo sapiens.cDNA clone
IMAGE:293827 5' similar to contains Alu repetitive element;; mRNA
sequence.
ACCESSION N98285
VERSION N98285.1 GI:1269708
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 437)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
,R., Williamson,A., Wohlmann,P. and Wilson,R.
The WashU-Merck EST Project
Unpublished (1995)
CONTACT: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: EtrPrimer
High quality sequence stop: 390.
FEATURES
Location/Qualifiers
1..437
/organism="Homo sapiens"
/db_xref="GDB:3801674"
/db_xref="taxon:9606"
/clone="IMAGE:293827"
/clone_lib="Soares fetal liver spleen lNFLS"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: p7T3D (Pharmacia)
with a modified polylinker; Site.1: Pac I; Site.2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
[5' AACTGGAAGATTAATTAAGATCTTTTATTTTATTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified p7T3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 91 a 123 c 105 g 117 t
ORIGIN
Query Match 13.0%; Score 262; DB 10; Length 437;
Best Local Similarity 100.0%; Pred. No. 9.8e-69;
Matches 262; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1185 agtctcaactgtgtcccggtgagtgagtgagtgagtgagtgagtgagtgagtgagtc 1244
|||||
Db 1 AGTCTCACTCTGTGTCCCGGCTGGAGTGCAGTGGACAAATCTCGGCTCATTTGCAACCTC 60

```

```

Qy 1245 cgctcccggttcacagtattctcttgctcagctcccaagtaactgatatattacaggc 1304
|||||
Db 61 CGCTCCGCGGTTCAAGTGATTTCTTGCTCAGCTCCCAAGTAACTGATATACAGGC 120
|||||
Qy 1305 gccagccaccacacccccgctgattttttagtagtagagcgggttttccaccgt 1364
|||||
Db 121 GCCAGCCACACACCCCGCTGATTTTGTATTTTAGTAGAGACGGGGTTTCCCCAGT 180
|||||
Qy 1365 tgcgggggtggtctcaaacctttgacctcaagtgaaccacccgctgtgctcccaag 1424
|||||
Db 181 TGGCGGGCTGGTCTCAAACTCTTGACCTCAAGTGAACCAACCCCGCTGTGCCCTCCCAAG 240
|||||
Qy 1425 tgctggaattaccagcgtgagc 1446
|||||
Db 241 TGCTGGAATTTACCAGCGTGAGC 262
|||||

RESULT 11
N20300/c
LOCUS
DEFINITION
N20300 472 bp mRNA linear EST 18-DEC-1995
YX43b10.s1 Soares melanocyte 2NBHM Homo sapiens cDNA clone
IMAGE:264475 3', mRNA sequence.
ACCESSION N20300
VERSION N20300.1 GI:1125255
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 472)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
,R., Williamson,A., Wohlmann,P. and Wilson,R.
The WashU-Merck EST Project
Unpublished (1995)
CONTACT: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 344
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 541 Std Error: 0.00
Seq primer: ml3 -40 forward
High quality sequence stop: 344.
FEATURES
Location/Qualifiers
1..472
/organism="Homo sapiens"
/db_xref="GDB:3874117"
/db_xref="taxon:9606"
/clone="IMAGE:264475"
/clone_lib="Soares melanocyte 2NBHM"
/sex="Male"
/tissue_type="melanocyte"
/lab_host="DH10B (ampicillin resistant)"
/note="Vector: p7T3D (Pharmacia) with a modified
polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TGTACCAATCTGAAGTGGAGCGCGCGCTTTTATTTTATTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified p7T3 vector
(Pharmacia). Library constructed by Bento Soares and
M.Fatima Bonaldo. RNA from normal foreskin melanocytes
(FS374) was kindly provided by Dr. Anthony P. Albino."
BASE COUNT 148 a 74 c 79 g 170 t
ORIGIN

```

Query Match 12.6%; Score 254; DB 10; Length 472;
 Best Local Similarity 100.0%; Pred. No. 2.2e-66;
 Matches 254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1753 gcatttaattatttcttctgacatttctgcaagcttggatttatatttccacttt 1812
 |||||
 Db 350 GCATTAAATTATTATTTCTGACATTTCTGCAAGCTTTGTATTATATTTCCACTTT 291
 |||||

QY 1813 atagatgggaatttgagctcttagaggtataaatactgcccaggtcacacagaag 1872
 |||||
 Db 290 ATAGATGAGGAATTTGAGGCTCTTAGAGGTAAATGACTTCCAGGTCCACAGGAAG 231
 |||||

QY 1873 tggcagacagctcttttaataagaataaataataataataataatgagataactt 1932
 |||||
 Db 230 TGGCAGACAGAGCTTTTAAATAAGAAAAAATAATAAATAATATGAGACTAACTT 171
 |||||

QY 1933 aaaataataaaccacaatttttaataataataataacgctgatacccaactaataaag 1992
 |||||
 Db 170 AAAATATTAAATACCACAAATTTTAAATTAATTAACCGTGTATAACCAACATTAATAAAG 111
 |||||

QY 1993 ttaagataccaaaa 2006
 |||||
 Db 110 TTAGATACCAAAA 97
 |||||

RESULT 12
 LOCUS BG614513 1066 bp mRNA linear EST 18-APR-2001
 DEFINITION 602642663F1 NIH_MGC_61 Homo sapiens cDNA clone IMAGE:4773711 5',
 mRNA sequence.

ACCESSION BG614513
 VERSION BG614513.1 GI:13665884
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1066)
 NIH-MGC <http://mgi.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: CLONTECH Laboratories, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L1CMI645 row: n column: 16
 High quality sequence stop: 613.
 Location/Qualifiers

FEATURES
 source
 1..1066
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4773711"
 /clone_lib="NIH_MGC_61"
 /tissue_type="embryonal carcinoma"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1:
 SfiI (ggccgctcgcc); Site_2: SfiI (ggccattatggcc);
 Double-stranded cDNA was prepared from cell line RNA. 5'
 and 3' adaptors were used in cloning as follows: 5'
 adaptor sequence: 5'-CACGGCCATTTAGCC-3' and 3' adaptor
 sequence: 5'-ATTCTAGAGCGGCGGCGGACATG-DT(30)BN-3'
 (where B = A, C, G and N = A, C, G, or T). Average
 insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies
 contained inserts by PCR. This library was enriched for
 full-length clones and was constructed by Clontech
 Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
 Library."
 307 a 272 c 240 g 247 t

ORIGIN

Query Match 9.6%; Score 194; DB 10; Length 1066;
 Best Local Similarity 100.0%; Pred. No. 9e-49;
 Matches 194; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 208 agcactactcccatgctccagctgtgacggaggttccacatcatattccagaag 267
 |||||
 Db 118 AGCATACTTCCCATTCAGCTGTTCACGGAGGTTTCACATCATATTTCCAGAAG 177
 |||||

QY 268 gctcctggaaagagtgtaatatgtgtcgcatccagagagctgattgggattgtgacttggc 327
 |||||
 Db 178 GCTCCTGGAAGAGTGAATATGTGTCATCCAGAGAGCTGATGGGATTTGACTTGGC 237
 |||||

QY 328 tgcgtcatctcttcattgctcaagcgcagaagaatctgtgcagccgcacacataactgt 387
 |||||
 Db 238 TGTGTGTCATCTTCATGTCAAGCGCAGAGAATCTGTGTGTCAGCCGCCACACATACTGT 297
 |||||

QY 388 taagcagtggtatga 401
 |||||
 Db 298 TAAGCAGTGGATGA 311
 |||||

RESULT 13
 LOCUS N27569 487 bp mRNA linear EST 30-DEC-1995
 DEFINITION YX43b10.r1 Soares melanocyte 2NbHM Homo sapiens cDNA clone
 IMAGE:264475 5', mRNA sequence.

ACCESSION N27569
 VERSION N27569.1 GI:1142050
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 487)
 Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman
 M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,
 Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston
 R., Williamson, A., Wohldmann, P. and Wilson, R.
 The WashU-Merck EST Project
 Unpublished (1995)
 Contact: Wilton RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 High quality sequence stops: 303
 Source: IMAGE Consortium, LLNL
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 541 Std Error: 0.00
 Seq primer: T7
 High quality sequence stop: 303.
 Location/Qualifiers

FEATURES
 source
 1..487
 /organism="Homo sapiens"
 /db_xref="GDB:3874117"
 /db_xref="taxon:9606"
 /clone="IMAGE:264475"
 /clone_lib="Soares melanocyte 2NbHM"
 /sex="Male"
 /tissue_type="melanocyte"
 /lab_host="DH10B (ampicillin resistant)"
 /note="Vector: pT73D (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5',
 TGTACCATCTGAAGTGGAGCGGCGGACGATTTTTTTTTTTT 3'],
 double-stranded cDNA was size selected, ligated to Eco RI
 adapters (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified pT73 vector

(Pharmacia). Library constructed by Bento Soares and M.Patima Bonaldo. RNA from normal foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P. Albino.

BASE COUNT 114 a 112 c 109 g 147 t 5 others

ORIGIN

Query Match 9.5% Score 191; DB 10; Length 487;
Best Local Similarity 99.6%; Pred. No. 1.4e-47;
Matches 241; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1447 caccatgcccgggtcacacgtttgagttgataccattgtgcattctcttttggcctct 1506
|||||
Db 102 CACCATGCGGGCTCACACGGTTTCAGTTGATACCATTTGTCATTCTCTTTGGCCTCT 161
|||||

QY 1507 tttttgctcatagagggttcaagataagtagtgtaagagccagtagtggtcataagaagc 1566
|||||
Db 162 TTTTGTGTCATAGAGGCTTCAAGATAGATTAGGTAAAGAGCCAGTAGTGTTCATAGAAGC 221
|||||

QY 1567 caatagagagcaggagccactttatcagtggtgaggtgtccgggctccctgctgcta 1636
|||||
Db 222 CAATAGAGAGCAGGAGCCACTTTATCAGGTGGCAGGTGTCTGGGCTCTCCCTGCTGCTA 281
|||||

QY 1627 gtcccaagcgtggtgtggtccagatctcttgagggtgataatgggacacacagagcac 1686
|||||
Db 282 GTCCCAAGCGGTGTGTTGCCAGGATCTTTGGAGGTGATATGGACACACAGAGGCAC 341
|||||

QY 1687 tg 1688
||

Db 342 TG 343

RESULT 14
BG532385
LOCUS 602561892F1 NIH_MGC_61 Homo sapiens cDNA clone IMAGE:469822 5',
DEFINITION mRNA sequence.
ACCESSION BG532385
VERSION BG532385.1 GI:13523923
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 798)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-femail.nih.gov
Tissue Procurement: AFCC
CDNA Library Preparation: CLONETECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLC1532 row: g column: 23
High quality sequence stop: 729.

FEATURES
source
1. .798
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:469822"
/clone_lib="NIH_MGC_61"
/tissue_type="embryonal carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1: SfII (ggcgcctcgcc); Site_2: SfII (ggcattatggcc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGCGGCGGCACATG-DT(30)BN-3'

(where B = A, C, or G and N = A, C, G, or T). Average insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."

BASE COUNT 240 a 185 c 193 g 180 t

ORIGIN

Query Match 7.1% Score 144; DB 10; Length 798;
Best Local Similarity 100.0%; Pred. No. 1e-33;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 439 ttgccacaggaagaacacacatggcaagaggaacagtaaacaggcacatcacagggaaca 498
|||||
Db 326 TTGCCACAGGAAGAACACCATGGCAAGAGGAGCAAGTAAACAGGCACATCAGGGGAACA 385
|||||

QY 499 cgaacatacaggccataaaactcttattagagagtcctacagataaatctctacagagcaa 558
|||||
Db 386 CGAAACATACGGCCATAAAACTCTTTATTAGAGAGTCTACAGATAAATCTACAGAGCAA 445
|||||

QY 559 ttctcaagtggacttgccatga 582
|||||
Db 446 TTCCTCAAGTGGACTTGGCCATGA 459
|||||

RESULT 15
AW083576/c
LOCUS xc18g08.x1 NCI_CGAP_Col9 Homo sapiens cDNA clone IMAGE:2584670 3',
DEFINITION mRNA sequence.
ACCESSION AW083576
VERSION AW083576.1 GI:6038728
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 294)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-femail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: Christa Prange, The I.M.A.G.E. Consortium DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -400P from Gibco
High quality sequence stop: 285.

FEATURES
source
1. .294
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2584670"
/clone_lib="NCI_CGAP_Col9"
/tissue_type="moderately differentiated adenocarcinoma"
/lab_host="DH10B"
/note="Organ: colon; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt. Normalized to Cot 50. Average insert size 1.32kb. Normalized version of NCI_CGAP_Col8. Library constructed by Life Technologies."
BASE COUNT 83 a 38 c 40 g 132 t 1 others
ORIGIN

Query Match 7.0%; Score 141; DB 9; Length 294;
Best Local Similarity 99.5%; Pred. No. 1.9e-32;
Matches 191; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1815 agatgagaaatttgaggctcttagaggtaaaatgactgcccaggtcacacaggaagtg 1874
|||||
DB 294 AGATGAGGAAATTTGAGGCTCTTAGAGGTAAATGACTTGCCAGGTCCACAGGAAGTG 235
|||||

QY 1875 gcagagacaagctttttaataaagaaaaaattataataataatagagtaacttaa 1934
|||||
DB 234 GCAGAGACAAGCTTTTAAATAAGAAAAAATTATAAATATATATGAGAGTAACTTAA 175
|||||

QY 1935 aatattaataaccacaatttttaataattaaccgtgataaccaacattataaaaagtt 1994
|||||
DB 174 AATATTAAATAACCNCAATTTAAATTAAATTAACCGTGATAACCAACATTATAAAAAGTT 115
|||||

QY 1995 aagataccaaaa 2006
|||||
DB 114 AAGATACCAAAA 103
|||||

Search completed: October 7, 2002, 12:07:31
Job time: 9086 sec

Pf		12-JAN-1999;	99WQ-USO0651.	
Xx		09-JUL-1998;	98US-0092155.	
Pr		20-JAN-1998;	98US-0071899.	
Xx		(CODO-) CODON DIAGNOSTICS LLC.		
Pa	Dyster LM,	Frustaci JM,	Papsidero LD;	
Xx				
Pt	WTPI:	1999-458469/38.		
Dr	P-PSDB:	AAY29092; AAY29093.		
Xx				
Ps	Claim 19;	Page 18-20;	76pp; English.	
Xx	The invention provides an isolated human chemokine, which is preferentially expressed in breast tissue or detected in breast milk. An antibody that recognizes the novel chemokine, or a chemokine-derived antigenic peptide, can be used to treat breast disease in a patient. A peptide, which binds to a cellular receptor for the chemokine, can also be used to treat breast disease. Antigenic peptides of the chemokine can be used to vaccinate patients against breast disease. The chemokine polynucleotide sequences and the chemokine protein can be detected in samples with primers, probes and antibodies using standard techniques. This is useful for detecting breast disease. Other breast diseases that may be treated or detected with the chemokine and its encoding polynucleotides include inflammations, infections, mastitis, benign cystitis, and benign hyperplasias as well as other malignancies. The present sequence represents the nucleotide sequence of the human mammary associated chemokine (MACK).			
Xx				
Sq	Sequence 3117 BP;	860 A;	582 C;	566 G;
				921 T;
				188 other;
	Query Match	51.3%;	Score 1035.2;	DB 20;
	Best Local Similarity	86.5%;	Pred. No. 5.7e-206;	Length 3117;
	Matches 1172;	Conservative 73;	Mismatches 92;	Indels 18;
	Gaps 1;			
Qy	671 ggtatgcaaatgatgccataataatactcacaactcttggtccaagcgatccccaccc	730		
Db				
	1669 ggctcatagtggcccagggttgattcctaaccctctgggtccaagcgatccccaccc	1728		
Qy	731 tagctcccgaagtagctgatataggttgtagccacagtcgcctggcataatttttc	790		
Db				
	1729 tagctcccgaagaacctggttatagggtgtagcacagtcgctggcataatttttc	1788		
Qy	791 ttgtgatcaaattcagggttaaaagtgttttgtttaagaattcttcacgtgaattcgtgtac	850		
Db				
	1789 ttgtgatcaaattcagggttaaagtgttttgtttaagaattcttcacgtgaattcgtgtac	1848		
Qy	851 tttatttgtcatttagagtttccataaatattagggtttatttttccataatagaatagtttaa	910		
Db				
	1849 tttatttgtcatttagagtttccataaatattagggtttatttttccataatagaatagtttaa	1908		
Qy	911 actaaaataaacttcaaaaaagtcgtctatttgcacagAACCTTTGCTCCGCCCGT	970		
Db				
	1909 actaaaataaacttcaaaaaagtcgtctatttgcacagAACCTTTGCTGGATGAATAATT	1968		
Qy	971 ctgatactgAAAAGACA AAAAAGCTGCCTTCCTGCCAGAACCTTTTGCTCCGCCCGT	1030		
Db				
	1969 ctgatactgAAAAGACA AAAAAGCTGCCTTCCTGCCAGAAACCSNTTGCTCCCCCAGT	2028		
Qy	1031 cagttcttgagcagcactagtttaggggccagagttcgccctctctggttggtgatttca	1090		
Db				
	2029 nagttcttgngcagnactagtttagggncocagagttngcccttnngkgtggfgatttca	2088		
Qy	1091 cgctctccctaaaacaaagagccttacatcttttagctcctattccacaccttctcacagtt	1150		
Db	: :	:		
	2089 ngvtctccctaaaacaaagngcnwacattytttagctcctattccacycttctnmacmamt	2148		
Qy	1151 tttatttgtttgtttgttt	1209		
	t-t-g-a-c-a-g-a-g-t-c-t-c-a-c-t-g-t-t-g-c-c-a-g-g-c-t-q-			

CC consensus derived from several overlapping and/or extended cDNA clones.
CC GFRP-2 is probably a CC chemokine and has chemical and structural
CC homology with hRECK (20% identity). GFRP-2 and hRECK are both basic
CC proteins, having isoelectric points of 10.1 and 10.2, respectively.
CC GFRP-2 was found by Northern analysis to be expressed in both tumorous
CC and nontumorous breast tissue. GFRP proteins (AAB03000-B03003),
CC nucleotides encoding them (AAS2455-A52458), GFRP agonists and
CC antagonists may be used to treat a wide variety of diseases associated
CC with increased or decreased expression or activity of GFRP proteins.
CC Conditions which may be treated include developmental disorders, cell
CC proliferative disorders (e.g., cancers), immune disorders (e.g.,
CC allergies, asthma), reproductive disorders (e.g., menstrual cycle
CC disorders), cardiovascular disorders (e.g., arteriosclerosis) and
CC bacterial, viral, fungal or parasitic infections. Additionally, GFRP
CC proteins and nucleotides can be used in the diagnosis of such disorders.
XX
SQ Sequence 566 BP; 152 A; 137 C; 144 G; 133 T; 0 other;

Query Match 21.9%; Score 441; DB 21; Length 566;
Best Local Similarity 96.0%; Pred. No. 1.5e-82;
Matches 485; Conservative 0; Mismatches 15; Indels 5; Gaps 3;

QY 97 tcgaacagcctcacttgttctgtcagtcagtaggcaggcagggaatgcagcagag 156
Db 1 tcgaacagcctcacttgttctgtcagtcagtaggcaggcagggaatgcagcagag 60
QY 157 aggaactgcgcattgtgcttgcctgtctgtgcggccctacatgcctcagaagccatact 216
Db 61 aggaactgcgcattgtgcttgcctgtctgtgcggccctacatgcctcagaagccatact 120
QY 217 tccattgctccagctgttcacgaggtttcacatcatatttccagagggctcctgga 276
Db 121 tccattgctccagctgttcacgaggtttcacatcatatttccagagggctcctgga 180
QY 277 aagagtgaatatgttcgcataccagagagctgattggggattgtgacttgctgctgtcat 336
Db 181 aagagtgaatatgttcgcataccagagagctgattggggattgtgacttgctgctgtcat 240
QY 337 ccttcattgccaagcagcagaagaatctgttcagcccgcaaccatactgttaagcagtg 396
Db 241 ccttcattgccaagcagcagaagaatctgttcagcccgcaaccatactgttaagcagtg 300
QY 397 gatgaagtgcgaagctgcgaagaaatgttaagaaatgtttgcacaggaagaacaa 456
Db 301 gatgaagtgcgaagctgcgaagaaatgttaagaaatgtttgcacaggaagaacaa 360
QY 457 ccattggcaagaggaacatacaggggcacatcagggggaacacgaacatacggccataa 516
Db 361 ccattggcaagagggacagtaacaggggcacatcagggggaacacgaacatacggccataa 420
QY 517 aactccttatta-gagagctcacagataaact--tacagagacaattcctcagtggaact 573
Db 421 aactccttattaggagagctaccggtaaatcttcocagagaccattcctcagtggaact 480
QY 574 t--ggccatgattggttgaagttt 596
Db 481 ttggccctgattggttgaagttt 505

RESULT 8
AAS76558/c
ID AAS76558 standard; cDNA; 675 BP.
XX
XX AAS76558;
AC
XX

DT 13-FEB-2002 (first entry)
XX
XX DNA encoding novel human diagnostic protein #12362.
DE
XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX

OS Homo sapiens.
XX
PN WO2001175067-A2.
XX
XX
PD 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.
PF 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
PA Drmanac RT, Liu C, Tang YT;
PI
XX
DR WPI: 2001-639362/73.
DR P-PSDB; ABG12371.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID No 12362; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 675 BP; 152 A; 166 C; 150 G; 207 T; 0 other;

Query Match 21.7%; Score 436.8; DB 23; Length 675;
Best Local Similarity 97.4%; Pred. No. 1.2e-81;
Matches 487; Conservative 0; Mismatches 7; Indels 6; Gaps 4;

QY 208 agccatacttccctcagctgttcacgaggtttcacatcatatttccagaag 267
Db 528 agccatacttccctcagctgttcacgaggtttcacatcatatttccagaag 469
QY 268 gctcct-ggaagagtgatattgtcgcataccagagagctgattgagctt-g 325
Db 468 gctcctgggaaagagtgatattgtcgcataccagagagctgattgagctt-g 409
QY 326 gctcgtgctatccttcagctcaa---gcgcagaagaatctgtg-tcagccgcacaacca 381
Db 408 gctcgtgctatccttcagctcaa---gcgcagaagaatctgtg-tcagccgcacaacca 349
QY 382 tactgttaagcagtgatgaaagtgcgaagctgcgaagaaatggttaaggaaattgtt 441
Db 348 tactgttaagcagtgatgaaagtgcgaagctgcgaagaaatggttaaggaaattgtt 289
QY 442 ccacagggaagaacaccatggaagaggaacagtaacaggggcacatcagggggaacaca 501
Db 288 ccacagggaagaacaccatggaagaggaacagtaacaggggcacatcagggggaacaca 229

QY 207 aagccatactccattccctccagctgtgtcagcgaggtttccatcatcatattccagaa 266
Db 61 aagccatactccattccctccagctgtgtcagcgaggtttccatcatcatattccagaa 120
QY 267 ggcctctggaagagtgatgtgtcgcacccagagagctgatgggattgtgacttgg 326
Db 121 ggcctctggaagagtgatgtgtcgcacccagagagctgatgggattgtgacttgg 180
QY 327 ctgctgtcatcttcatgtcaagcgacagaagaatctgtcagccgcacacacatactg 386
Db 181 ctgctgtcatcttcatgtcaagcgacagaagaatctgtcagccgcacacacatactg 240
QY 387 ttaagcagtgatgaaagtcaagctgccaagaaaaatgtaaaaggaaatgtttgccaca 446
Db 241 ttaagcagtgatgaaagtcaagctgccaagaaaaatgtaaaaggaaatgtttgccaca 300
QY 447 ggaagaacacacattggcgaaggaacagtaacacaggggcacacacacacacacac 506
Db 301 ggaagaacacacattggcgaaggaacagtaacacaggggcacacacacacacacac 360
QY 507 acggccataaaactcttatt 527
Db 361 acggccataaaactcttatt 381

RESULT 12

AAV38292
ID AAV38292 standard; cDNA; 496 BP.

XX AC AAV38292;

XX DT 12-OCT-1998 (first entry)

XX DE Human C-C chemokine DVic-1 EST HHFFQ25R.

XX KW DVic-1; DNAX Vic-1; C-C chemokine; cytokine; human; immune system;
XX KW cancer; cell proliferation; therapy; diagnosis; HHFFQ25R; ss.

XX OS Homo sapiens.

XX PN WO9823750-A2.

XX PD 04-JUN-1998.

XX PF 26-NOV-1997; 97WO-US21092.

XX PR 05-DEC-1996; 96US-0761071.

XX PR 27-NOV-1996; 96US-0031805.

XX PA (SCHE) SCHERING CORP.

XX PI Hedrick JA, Morales J, Vicari A, Zlotnik A;

XX DR WPI; 1998-322730/28.

XX PT DVic-1 and DGWCC chemokines - useful for developing products for
XX PT treating abnormal physiology or development, e.g. cancerous or
XX PT degenerative conditions

XX PS Disclosure; Page 61; 71pp; English.

XX CC This sequence defines EST HHFFQ25R obtained from a human foetal
XX CC heart library. It shows high homology to EST HOEDH1LR (see AAV38293)
XX CC obtained from a human osteoblast library. The two ESTs are
XX CC probably from a similar transcript. The chemokine motifs of the
XX CC two ESTs were compared, and a consensus sequence was derived and
XX CC subsequently confirmed as encoding human DNAX Vic-1 (DVic-1) (see
XX CC AAW60649), a novel C-C chemokine. cDNA (see AAV38291) encoding DVic-1
XX CC can be obtained from e.g. skin, epithelial or wound healing
XX CC libraries. DVic-1 plays a role in the regulation or development of
XX CC neuronal or haematopoietic cells, e.g. lymphoid cells, which affect
XX CC immunological responses. It can be used in the treatment of
XX CC conditions associated with abnormal physiology or development,

CC including abnormal proliferation, e.g. cancerous conditions or
CC degenerative conditions. Abnormal proliferation, regeneration,
CC degeneration, and atrophy may be modulated by appropriate
CC therapeutic treatment using products of the invention. The products
CC can also be used for detection, diagnosis and drug screening.

XX SQ Sequence 496 BP; 119 A; 100 C; 133 G; 118 T; 26 other;

Query Match 18.3%; Score 369.6; DB 19; Length 496;

Best Local Similarity 85.0%; Pred. NO. 1e-67;

Matches 420; Conservative 0; Mismatches 70; Indels 4; Gaps 2;

QY 91 tctctgacgaacagctccactgtgtgtgtcagtgccagtagggcagcgagaaatgca 150
Db 1 tctctgacgaacagctccactgtgtgtgtcagtgccagtagggcagcgagaaatgca 60

QY 151 gcagagagagactgcacatctggccttggctgtctgtgcccctcatctcctcagaagc 210
Db 61 gcagagagagactgcacatctggccttggctgtctgtgcccctcatctcctcagaagc 120

QY 211 catacttccattgctccagctgtgtgacaggggtttcacatcatatttccagaagcct 270
Db 121 catacttccattgctccagctgtgtgacaggggtttcacatcatatttccagaagcct 180

QY 271 cct-ggaagagtgaaatgtgtcgcacccagagagctgtggtgattgtgacttgctg 329
Db 181 ccggggaagagtgaaatgtgtcgcacccagagagctgtggtgattgtgacttgctg 240

QY 330 ctgtcatcttcatgtcaagcgacagaagaatctgtgtcagccgcacacacacacacac 389
Db 241 ctgtcatcttcatgtcaagcgacagaagaatctgtgtcagccgcacacacacacacac 300

QY 390 agcagtgagtgaaagtcaagctgccaagaaaaatgtaaaaggaaatgttgcacagga 449
Db 301 agcagtgagtgaaagtcaagctgccaagaaaaatgtaaaaggaaatgttgcacagga 360

QY 450 ag---aaacacacatgccaaggaagacagtaacacaggggcacacacacacacacac 506
Db 361 ngaaacacacacatgccaaggaagacagtaacacaggggcacacacacacacacac 420

QY 507 acggccataaaactcttatttagagagcttacagataaaatctacagagacaattctc 566
Db 421 ngggcctnaaaatctctttnngggntttaagggttaaaattttnnnggaaattttccna 480

QY 567 gtggacttgccat 580

Db 481 ggggntttggncat 494

RESULT 13

AAV89394/C

ID AAV89394 standard; DNA; 311 BP.

XX AC AAV89394;

XX DT 29-SEP-1999 (first entry)

XX DE Human MACK hybridising antisense riboprobe.

XX KW Chemokine; breast tissue; breast milk; breast disease; vaccine; human;
XX KW inflammation; infection; mastitis; benign cystitis; hyperplasia;
XX KW mammary associated chemokine; MACK; probe; ss.

XX OS Synthetic.

XX OS Homo sapiens.

XX PN WO936540-A1.

XX PD 22-JUL-1999.

XX PF 12-JAN-1999; 99WO-US00651.

XX XX

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 7, 2002, 09:01:55 ; Search time 1708.76 Seconds
(without alignments)
15931.645 Million cell updates/sec

Title: US-09-813-492-1
Perfect score: 2017
Sequence: 1 tagataccctgaacacccctcc.....atatcaaaaaaaaaaaaaa 2017

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST: *

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estmu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_htc:*
- 9: gb_estl:*
- 10: gb_est2:*
- 11: gb_htc:*
- 12: gb_gss:*
- 13: em_gss_hum:*
- 14: em_gss_inv:*
- 15: em_gss_pln:*
- 16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	734.2	36.4	816	10 BG682175	BG682175 602629474
2	638.2	31.6	698	10 BF034600	BF034600 601455024
c 3	530.2	26.3	577	10 N63913	N63913 za27h10.s1
c 4	508	25.2	508	9 AW469978	AW469978 xr27a11.x
5	497.8	24.7	869	10 BG530240	BG530240 602559348
6	454.2	22.5	798	10 BG532385	BG532385 602561892
c 7	448.4	22.2	450	9 A1880389	A1880389 at55a07.x
8	434	21.5	1066	10 BG614513	BG614513 602642663
9	418.6	20.8	873	10 BE748488	BE748488 601571978
10	389.2	19.3	437	10 N98285	N98285 za27h10.r1
c 11	377.2	18.7	399	9 BE088415	BE088415 CM0-BT068
c 12	361.8	17.9	472	10 N20300	N20300 yx43b10.s1
c 13	360.4	17.9	362	10 BE463561	BE463561 hw25b05.x
14	301	14.9	345	12 AZ575795	AZ575795 ASP-T29B0
c 15	292.6	14.5	510	9 BE004848	BE004848 MR2-BN011
16	283.4	14.1	487	10 N27569	N27569 yx43b10.r1
17	281.8	14.0	695	9 BB617469	BB617469 BB617469

18	280.2	13.9	683	9	BB661838
19	273.2	13.7	751	10	BF785656
20	243.2	12.1	714	10	BG249748
c 21	230.6	11.4	477	9	BE150796
c 22	230.2	11.4	673	10	BG493959
c 23	229.8	11.4	642	12	B59854
c 24	229	11.4	738	10	BG527374
c 25	228	11.3	643	10	BF212465
c 26	227.6	11.3	888	10	BF184128
c 27	227	11.3	687	10	BG777306
c 28	226.6	11.2	782	10	BG741134
c 29	226.2	11.2	417	10	R38459
c 30	226	11.2	416	9	BE062476
c 31	225.6	11.2	821	9	AL520459
c 32	225.4	11.2	412	9	BE062478
c 33	225.4	11.2	909	10	BF184981
c 34	224.4	11.1	638	10	BG500685
35	224.4	11.1	493	12	B50449
c 36	224.4	11.1	894	10	BG328312
c 37	223.8	11.1	694	10	BG776604
c 38	223.6	11.1	490	12	AQ124226
c 39	223.4	11.1	676	12	AG178621
c 40	223	11.1	521	9	AW970571
c 41	222.4	11.0	708	10	BF978936
c 42	222.4	11.0	841	10	BF663361
c 43	222	11.0	502	9	AI380617
c 44	221.8	11.0	1273	11	BC011712
45	221.4	11.0	397	10	BG386752

ALIGNMENTS

RESULT# 1
BG682175 816 bp mRNA linear EST 01-MAY-2001
LOCUS 602629474F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4754250 5',
DEFINITION mRNA sequence.
ACCESSION BG682175.1 GI:13913572
VERSION BG682175.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 816)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10615 row: c column: 19
High quality sequence stop: 766.

FEATURES
Location/Qualifiers
1..816
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4754250"
/clone.lib="NCI_CGAP_Skn4"
/tissue.type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: skin; Vector: PCMV-SPOPT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.5kb. Library constructed by Life Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 184 a 199 c 172 g 261 t

Qy	1830	aggctcttagaggttaaataactgacctggccagggtcacacaggaagtgcagagacaagcttt	1889
Db	541	AGGCTCTTAGAGGTAAAATGACTTGCCAGGTACACAGGAAGTGGCAGACAGAAGCTTT	600
Qy	1890	ttaataagaagaaaaataataaaataataatgatgagagtaacttaaaaataataaacca	1949
Db	601	TTAATAAAGAAAAAAATTTAAATATATATGAGAG---TACTTAAAAATATTAAATA--CA	655
Qy	1950	caattttaaatattaacacgtgataaccaacaacattaataaa	1990
Db	656	CATTTTTAATTAA-PAACCGTGATAACCAACATTTATATAA	695
RESULT	3		
N63913/c			
LOCUS	N63913	577 bp	mRNA linear EST 01-MAR-1996
DEFINITION	za27hl0.sl Soares fetal liver spleen INFLS Homo sapiens cDNA clone IMAGE:293827 3' similar to contains element MIR repetitive element ;, mRNA sequence.		
ACCESSION	N63913		
VERSION	N63913.1	GI:1211742	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 577)		
TITLE	Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston ,R., Williamson,A., Wohldmann,P. and Willson,R. The WashU-Merck EST Project Unpublished (1995) Contact: Wilson RK Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu		
JOURNAL	This clone is available royalty-free through LLNL : contact the IMAGE Consortium (info@image.llnl.gov) for further information. Seq primer: ml3 -40 forward High quality sequence stop: 356.		
COMMENT	Location/Qualifiers		
FEATURES	source		
	1..577	/organism="Homo sapiens"	
		/db_xref="GDB:3801674"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:293827"	
		/clone_lib="Soares fetal liver spleen INFLS"	
		/sex="male"	
		/dev_stage="20 week-post conception fetus"	
		/lab_host="DH10B (ampicillin resistant)"	
		/notes="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Pac I; Site_2: Eco RI; 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5' ACTCGACAATATTAAGATCTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."	
BASE COUNT	168 a	135 C	132 g
ORIGIN		141 t	1 others
Query Match	26.3%	Score 530.2;	DB 10; Length 577;
Best Local Similarity	98.8%;	Pred. No. 3.1e-69;	
Matches	565; Conservative	0; Mismatches	4; Indels 3; Gaps 3;
Qy	1329	tttgtatttttagtagagacggggttttccccagtggtgcggcggtgctcacaactctt	1388

/db_xref="taxon:9606"
/clone="IMAGE:2761340"
/clone_lib="NCI_CGAP_Ut4"
/tissue_type="serous papillary carcinoma, high grade, 2
pooled tumors"
/lab_host="DH10B"
/note="Organ: uterus; Vector: PCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.48 kb. Life Technologies catalog #:
11542-016"

BASE COUNT 150 a 102 c 92 g 164 t
ORIGIN

Query Match 25.28; Score 508; DB 9; Length 508;
Best Local Similarity 100.0%; Pred. No. 6.2e-66;
Matches 508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1495 cttttggcctctttttgtccatagagcgttcaagatagataggttaagagcccgagtagtg 1554
Db 508 CTTTTGGCCTCTTTTGTCCATAGAGCCTCAAGATAGATAGGTAAAGCCCAAGTAGTG 449
Qy 1555 ttctaagaagcccaatagagagcgagggccactttatcaggtggcaggtgtcccgggcct 1614
Db 448 TTTCAAGAGCCCAATAGAGAGCAGGAGCCACTTTATCAGGTGGCAGGTGTCCCGGGCCT 389
Qy 1615 ccctgctgctagtcaccaagcgggtgttgcagaggtctcttgaggtgataatgggac 1674
Db 388 CCCTGCTGGCTAGTCCCAAGCGGTGGTGTGCCAGGATGCTTTGAGGTGATAATGGGAC 329
Qy 1675 acacagagggcactgagctctccataggttaaaatgccacaaaactggccttggcctaata 1734
Db 328 ACACAGAGGCCTGAGTCTCCATAGGTAAATGCCACCAAACTGGCCTTGGCTTAATA 269
Qy 1735 tccctcattgactatttagcaatttaattttatttcttcagcattctgcaagctttg 1794
Db 268 TCCCTCATGACTATTAGCAATTAATTTATTTATTTCTCTGACATTTCTGCAAGCTTTG 209
Qy 1795 tatttatcttccactttatagatagagaaatttgaggtctcttagagtaaatgacttg 1854
Db 208 TATTATATTTCCTACTTTATAGATGAGAAATTTGAGGCTCTTAGAGGTAAATGACTTG 149
Qy 1855 cccaggtcacacaggaagtggcagagacaagcttttttaataagaagaaaaataataaat 1914
Db 148 CCCAGGTCACACAGGAAGTGGCAGACAGACAGCTTTTAAATAAGAAAAATTAATAAAT 89
Qy 1915 ataataagagtaactataaataatataaaccacaaatttttaataatgacgtgat 1974
Db 88 ATAATATGAGAGTAACCTAAATATTAATAAACCACAAATTTTAAATTAATTAACCGTGAT 29
Qy 1975 aaccaacattaataaagttaagatacc 2002
Db 28 AACCAACATTAAATAAAGTTAAGATACC 1

RESULT 5
BG530240 869 bp mRNA linear EST 03-APR-2001
LOCUS 602559348F1 NIH_MGC_61 Homo sapiens cDNA clone IMAGE:4697338 5',
DEFINITION mRNA sequence.
ACCESSION BG530240
VERSION BG530240.1 GI:13521777
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 869)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-i@mail.nih.gov

Tissue Procurement: ATCC
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LICM1525 row: p column: 11
High quality sequence stop: 748.
Location/Qualifiers
1. .869
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4697338"
/clone_lib="NIH_MGC_61"
/tissue_type="embryonal carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggcgctcgcc); Site_2: SfiI (ggccattatggcc);
Double-stranded cDNA was prepared from cell line RNA. 5'
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CACGCCCATATGGCC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGGCGGCGGCGGCATAG-3' (30)BN-3'
(where B = A, C, G, or T). A, C, G, or T). Average
insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
Library."

BASE COUNT 261 a 185 c 215 g 208 t
ORIGIN
Query Match 24.78; Score 497.8; DB 10; Length 869;
Best Local Similarity 99.6%; Pred. No. 1.6e-64;
Matches 499; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 201 cctcagaagccatactcccatctccctcagctgttcacgaggtttcacatcatattt 260
Db 87 CCTCACCAGCCATCTTCCCATTTGCCCTCCAGCTGTTCACGAGAGCTGATGGGATTTT 146
Qy 261 ccagaagcctcctggaagagtgatgatgtgcacatccagagagctgatgggattgtg 320
Db 147 CCAGAAGGCTCCTGGAAAGAGTGAATATGTGTCGATCCAGAGAGCTGATGGGATTTG 206
Qy 321 acttggtgtgtcactccttcactgtcaagcgagagaatctgtgtcagcccgacacac 380
Db 207 ACTTGGCTGTGTCTCTTCATGTCAGGCGCAGAAGAATCTGTGTCAGCCCGCACAC 266
Qy 381 atactgttaagcagtgatgaagtcaagctgcaagtgccaaagaaatggttaaggaatgtt 440
Db 267 ATACTGTTAAGCAGTGGTGAAGTGAAGTGCAGTCCAGAAATATGGTAAGGAATGTTT 326
Qy 441 gccacaggaagaaacacacatggcgaaggaagaaacagtaacagggcacatcaggggaaacag 500
Db 327 GCCACAGGAAGAACACCATGTCGCAAGAGGAACACAGTAACAGGCGACATCAGGGGAACAG 386
Qy 501 aaacatacggccataaaactcctctattagagagctcacagataaaatctacagagaatt 560
Db 387 AAACATACGGCCATAAAACTCCTTTATTAGAGAGTCTACAGATAAATCTACAGAGACAA 446
Qy 561 cctcaagtgaacttggcagtgattgttaagtattcatctgaattcctctattgta 620
Db 447 CCTCAAGTGGACTTGGCCATGATTGGTTGTAAGTTTATCATCTGAATCTCCTTATTGTA 506
Qy 621 gacaacagaacaaacaaaatttggtgttttaaaaaataaacaaattgtgcggtatgcaaa 680
Db 507 GACAACAGAACAAACAAATATTGGTTTTTAAAAAATGAACAATTTGCGGTTATGCAAA 566
Qy 681 ttagccaataataatactcaa 701
Db 567 TGTAGCCAATAATATACTCAA 587

```

RESULT 6
BG532385
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 798)
AUTHORS
NIH-MGC http://mgc.ncl.nih.gov/
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM1532 row: 9 column: 23
High quality sequence stop: 729.
FEATURES
source
1..798
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4699822"
/clone_lib="NIH_MGC_61"
/tissue_type="embryonal carcinoma"
/lab_host="DH10B (Ti phage-resistant)"
/notes="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggccattggcc); Site_2: SfiI (ggccattggcc);
Double-stranded cDNA was prepared from cell line RNA. 5'
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGCGGCGGCACATG-dt(30)BN-3'
(Where B = A, C, G and N = A, C, G, or T). Average
insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
Library."
BASE COUNT 240 a 185 c 193 g 180 t
ORIGIN
Query Match 22.5%; Score 454.2; DB 10; Length 798;
Best Local Similarity 97.8%; Pred. No. 4.6e-38;
Matches 492; Conservative 0; Mismatches 8; Indels 3; Gaps 3;
QY 201 cctcagaagcactactcccatgtccctccagctgtgtgcaggggttcacatcatattt 260
|||||
DB 88 CCTCACCAGGCATCTACTCCATGTCCTCCAGCTGTTCACGGAGGTTTTCATCATCATATT 147
|||||
QY 261 ccag-aaggtcctcggaagtgaaatgtgtcgcaatccagagagctgatgggattgt 319
|||||
DB 148 CCAGAAAGGTTCCGGGAAAGAGTGAATATGTGTGCGCATCCAGAGAGCTGATGGGATGTT 207
|||||
QY 320 gaattggctgtctcatccttcattgcaagcgagagaagaatctgtctcagccgcacac 379
|||||
DB 208 GACTTGGCTGCTGTCTATCTTCATGTCAAGCGCAGAGAAATCTGTGTACGCCGCGCAAC 267
|||||
QY 380 cactagttaagcagtgagtgcaagtgcaagtgcaagaaataatggttaagaaatgtt 439
|||||
DB 268 CATACTGTTAAGCAGTGATGAAGATGCAAGCTGCCAAGAAATGTTAAGGAATG-T 326
|||||
QY 440 tgccacagggaagaacaccatggcaagagggaacagtaacaggggcaccatcaggggaaacac 499
|||||
Db 327 TGCCACAGGAAGAAACACCATGTGCAAGAGGAACAGTAAACAGGCACATCAGGGGAACAC 386
|||||
QY 500 gaaacatacagccataaaactcctattatagagagctacagataaaatctacagagacaat 559
|||||
Db 387 GAAACATACGCCCAATAAACTCCTATTATAGAGAGCTCAGAGATAAATCTCAGAGACAAT 446
|||||
QY 560 tctcaagtggacttgccatgattgttgtaagtattatcatctgaattcctctattgt 619
|||||
Db 447 TCCTCAAGTGCATTTGCCCATGAGTGGTTGTAAGTTTATCATCTGAATTCCTTATCGT 506
|||||
QY 620 agacaacagacacacacacacacacacacacacacacacacacacacacacacacacac 678
|||||
Db 507 AGACAACAGACAAACAAATAATTTGGTTTACACAAATGAACAATTTGTCGGGTATGCA 566
|||||
QY 679 aatgtagccaataataatactacaa 701
|||||
Db 567 AATGTAGCCAATAATAATACTCAA 589
|||||
RESULT 7
AI880389/c
LOCUS
DEFINITION
at55a07.x1 Barstead colon HPLRB7 Homo sapiens cDNA clone
IMAGE:2375892 3', mRNA sequence.
ACCESSION
AI880389
VERSION
AI880389.1 GI:5554438
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 450)
AUTHORS
Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Krizman,D., Kucaba,T., Lacy,M., Le.N., Lennon,G., Marra,M., Martin
,J., Moore,B., Scheinberg,K., Steptoe,M., Tan,F., Theising,B.,
White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE
WashU-NCI human EST Project
JOURNAL
Unpublished (1997)
COMMENT
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through LLNL : contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -400P from Gibco.
FEATURES
source
1..450
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2375892"
/clone_lib="Barstead colon HPLRB7"
/sex="male"
/dev_stage="adult, age 25"
/lab_host="DH10B (phage resistant)"
/notes="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st
strand cDNA was primed with a Not I - oligo(df) primer [5'
TGTTACGAATCTGAAGTGGAGCGGCCCTTTTTTTTTTTTTTTTTTTTTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
[5', AATTCACTAGTAAT 3' and 5' ATTACTAGT 3'], digested
with Not I and cloned into the Not I and Eco RI sites of
the modified pT73 vector. Library constructed by Bob
Barstead."
BASE COUNT 132 a 88 c 80 g 150 t
ORIGIN
Query Match 22.2%; Score 448.4; DB 9; Length 450;
Best Local Similarity 99.8%; Pred. No. 4.2e-57;
Matches 449; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```
QY 1553 tgttcataagaagccaatagagcagagcagcactttatcagggtgacaggtgtcccggtc 1612
|||||
Db 450 TGTTCATAAGAAGCCAATAGAGCAGGAGCCACATTATCAGGTGCAGGTGTCTGGGC 391
|||||
QY 1613 ctccctcgtcgttagtcccaagcgtgtgttgcaggatgtcttggaggtgataatg 1672
|||||
Db 390 CTCCCTGCTGGCTAGTCCCAAGCGGTGGTGTGCCAGGATGTCTTGGAGGTGATAATGGG 331
|||||
QY 1673 acacacagaggcactgagctcctcataggttaaaatgcccacaaactggcctttgcctaa 1732
|||||
Db 330 ACACACAGAGGCACTGAGTCTCATAGGTAAATGCGCCACCAAACTGGCCTTGGCCTAA 271
|||||
QY 1733 tatccctcattgactattgacatttaattattttcttctgacattctgcaagctt 1792
|||||
Db 270 TATCCCTCATTGACATATTAGCATTTAAATTTATTTTCTTGACATTTCTGCAAGCTT 211
|||||
QY 1793 tgtatttatctccactttatagtagagaaatttgagcctcttagaggttaaaatgact 1852
|||||
Db 210 TGTATTATATTCTCCACTTTATAGATGAGGAAATTTGAGGCTTTAGAGGTAATAATGACT 151
|||||
QY 1853 tgcacaggtccacaggaagtggcagagcagccttttaataagaagaaaaattataaa 1912
|||||
Db 150 TGCCCAAGGTCACACAGGAAGTGGCAGAGCAAGCTTTTTTAAATAAGAAAAATTAATAA 91
|||||
QY 1913 atataatgagagtaacttaataattataaaccacaaattttaataattaaccgtg 1972
|||||
Db 90 ATATAATATGAGATGAACCTAAATATTAATAAACCAACAAATTTAAATTAATTAACCGTG 31
|||||
QY 1973 ataacaacattataaaagttaagatacc 2002
|||||
Db 30 ATAACCAACATTAAATAAAGTTTAAGATACC 1
|||||

RESULT 8
BG614513.1 BG614513 1066 bp mRNA linear EST 18-APR-2001
LOCUS 602642663F1 NIH_MGC_61 Homo sapiens cDNA clone IMAGE:4773711 5',
DEFINITION mRNA sequence.
ACCESSION BG614513
VERSION BG614513.1 GI:13665884
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL NIH-MGC http://mgc.nci.nih.gov/
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONETECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM1645 row: n column: 16
High quality sequence stop: 613.
Location/Qualifiers
1..1066
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4773711"
/clone_lib="NIH_MGC_61"
/tissue_type="embryonal carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggcgcctcgccc); Site_2: SfiI (ggcctattggcc);
Double stranded cDNA was prepared from cell line RNA.
and 3' adaptors were used in cloning as follows: 5'
```

adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor
sequence: 5'-ATTAGAGCGCGGCGGCACATG-dT(30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
Library."

BASE COUNT 307 a 272 c 240 g 247 t

ORIGIN

Query Match 21.5%; Score 434; DB 10; Length 1066;

Best Local Similarity 95.1%; Pred. No. 3.9e-55;

Matches 481; Conservative 0; Mismatches 20; Indels 5; Gaps 3;

QY 201 cctcagaagcactactcctccattgctccagctgttgacgaggtttcacatcatattt 260

|||||

Db 111 CCTCACCAGCCACTACTTCCCATTTGCCAGCTGTTCACGGAGGTTTCACATCATATT 170

QY 261 ccagaaggtcctcgaagaagtgaaatatgtcgcatccagagagctgatgggattgtg 320

|||||

Db 171 CCAGAAGGCTCTGTGAAGAAGAGTAATATGTGCGCATCCAGAGAGCTGATGGGATTGTG 230

QY 321 acttgctgtgtcattccttcattgcaagcgcagagaatctgtgcagccgcgcacacc 380

|||||

Db 231 ACTTGGCTGCTGTCATCTTCAATGTCAAGCGCAGAGAATCTGTCTCAGCCCGCACAA 290

QY 381 atactgttaagcagtggtgaaagtcaagctgccaagaaaaatggtaaa--ggaaatgt 438

|||||

Db 291 ATACTGTTAAGCAGTGGATGACAGTGCAAGCTGCCAAGAAAAATGTTACACGGAACATGT 350

QY 439 ttgcccacaggaagaacaccatggcaagaggaacagtaaacaggggcacatcagggggaaca 498

|||||

Db 351 TTGCCACAGGAAGAAACACCATTTGGCAAGAGGAACAGTAACACGGGCACATCAGGGGGAACA 410

QY 499 cgaacatacgccataaaactcctatttagagagctctacagataaactacagagacaa 558

|||||

Db 411 CGAACAATACGCGCCATAAAACTCCTTATTAGAGAGCTTACAGATAAATCTACAGAGACAA 470

QY 559 ttctcaagtggaacttgccatgattggtgttaa--gtttatcatctgaattctcctatt 617

|||||

Db 471 TTCCCTCAAGTGACCTTGGCCATGATCGGTGTAAAGCTTTATCATCTGAATTTCTCCTATT 530

QY 618 gtagaacacagacaaaacaaaataattggtttt--taaaaaatgaacaattgtcggtat 675

|||||

Db 531 GTAGACACACCAAGAACAAACAAATATTGTTCTTCACACACAAAGAACAACTGTGCGGTAT 590

QY 676 gcaaatgtagccaataataatactcaa 701

|||||

Db 591 GCAATGTAGCCAATAATATACTCAA 616

RESULT 9

LOCUS BE748488

DEFINITION 601571978F1 NIH_MGC_55 Homo sapiens cDNA clone IMAGE:3838987 5',

ACCESSION BE748488

VERSION BE748488.1 GI:10162480

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

1 (bases 1 to 873)

AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

JOURNAL NIH-MGC http://mgc.nci.nih.gov/.

COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: CLONETECH Laboratories, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLCM526 row: k column: 20
 High quality sequence stop: 150.

FEATURES

Location/Qualifiers

1. 873

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:3838987"

/clone_lib="NIH_MGC_55"

/tissue_type="from acute myelogenous leukemia"

/lab_host="DH10B (T1 phage-resistant)"

/note="organ: bone marrow; Vector: pDNR-LIB (Clontech);

Site_1: SflI (ggcgctcgcc); Site_2: SflI (ggcattatggcc

); Double-stranded cDNA was prepared from cell line RNA.

5' and 3' adaptors were used in cloning as follows: 5'

adaptor sequence: 5'-CAGCGCATATGGCC-3' and 3' adaptor

sequence: 5'-ATTCTAGAGCGGAGCGCGCATG-dT(30)BN-3'

(where B = A, C, or G and N = A, C, G, or T). Average

insert size 1.65 kb (range 0.9-4.0 kb). 14/15 colonies

contained inserts by PCR. This library was enriched for

full-length clones and was constructed by Clontech

Laboratories (Palo Alto, CA)."

254 a 201 c 214 g 204 t

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches 474; Conservative

0; Mismatches 14; Indels 7; Gaps 4;

QY 201 cctcagaagccatacttccattgctccagctgtgtgcagaggtttcacatcatattt 260

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 66 cctcaccagccctacttccattgctccagctgtgtgcagaggtttcacatcatattt 125

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 261 ccagaaggtctctgaaagtgaaatgtgtcgcacatccagagagctgatgggattgtg 320

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 126 ccagaaggtctctgaaagtgaaatgtgtcgcacatccagagagctgatgggattgtg 185

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 321 atttggtctgtctatcttcattgtcgaagcgagaagaatctgtgtcagccgcacaaac 380

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 186 acttggtctgtctatcttcattgtcgaagcgagaagaatctgtgtcagccgcacaaac 245

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 381 atactgttaagcagtgatgaaagtgcgaagctgcgaagaaataatggtaaggaatgttt 440

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 246 atactgtttaagcagtgatgaaagtgcgaagctgcgaagaaataatggtaaggaatgttt 305

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 441 gccacaggaagaacaccatggcaagaggaacagtaacacagggcacatcaggggaacacag 500

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 306 gccacaggaagaacaccatggcaagaggaacagtaacacagggcacatcaggggaacacag 365

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 501 aacatacagggccataaaactccttattagagagctcacagataaaatctcacagagacaatt 560

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 366 aaacatacagggccataaaactccttattagagagctcacagataaaatctcacagagacaatt 425

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 561 cctcaagtgaattggccatgatt--ggttgtaagtattatcatctgaattctccttattg 618

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 426 cctcaagtgaattggccatgatt--ggttgtaagtattatcatctgaattctccttattg 484

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 619 tagacaacagaacaaataattgttttttaaaaaa---tgaacaattgtg-cggta 674

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 485 gagacaacagaacaaataattgttttttaaaaaa---tgaacaattgtg-cggta 544

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 675 tgcaaatgtagccaa 689

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 545 tgcaaatgttagccaa 559

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

RESULT 10

N98285

LOCUS

DEFINITION

Accession

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

N98285 437 bp mRNA linear EST 10-APR-1996
 za27h10.r1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone
 IMAGE:293827 5' similar to contains Alu repetitive element;; mRNA
 sequence.

N98285
 N98285.1 GI:1269708
 EST.
 human.

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 437)
 Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman,
 M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,
 Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston,
 R., Williamson, A., Wohlmann, P., and Wilson, R.

The WashU-Merck EST Project
 Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu

This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: ETPRimer
 High quality sequence stop: 390.

FEATURES

source

Location/Qualifiers

1. 437

/organism="Homo sapiens"

/db_xref="GDB:3801674"

/db_xref="taxon:9606"

/clone="IMAGE:293827"

/clone_lib="Soares fetal liver spleen INFLS"

/sex="male"

/dev_stage="20 week-post conception fetus"

/lab_host="DH10B (ampicillin resistant)"

/note="Organ: Liver and Spleen; Vector: p7T3D (Pharmacia)

with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;

1st strand cDNA was primed with a Pac I - oligo(dT) primer

[5' AACTGGAAGAAATAATAAGATCTTTTTTTTTTTTTTTT 3'],

double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Pac I and cloned into the Pac I

and Eco RI sites of the modified p7T3 vector. Library

went through one round of normalization. Library

constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 91 a 123 c 105 g 117 t

ORIGIN

Query Match 19.3%; Score 389.2; DB 10; Length 437;
 Best Local Similarity 97.3%; Pred. No. 2.4e-48;
 Matches 427; Conservative 0; Mismatches 9; Indels 3; Gaps 3;

QY 1185 agtctcactctgtgccaggtgagtgagtcagtgagcagacatctcgctcattgcaacctc 1244

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 1 AGTCTCACACTGTGTTGCCAGGCTGGAGTGCAGTGGCACAACTCTCGCTCATTCGCAACCTC 60

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 1245 cgctcccgcttcaagtgtattcttgcctcagcctcccaagttaactgatatcaggc 1304

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 61 CGCCTCCCGGTTCAAGTGATCTCTTGCTCAGCCTCCCAAGTAAGTATATACAGGC 120

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 1305 gccacgccaccacacccccgctgatttttatttttagtagagacggggttttcccaagt 1364

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 121 GCCCAGCCACACACCCCGCTGATTTTGTATTTTAGTAGAGACGGGGTTTCCACAGT 180

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 1365 tggcggggtgtctcaactcttgacctcaagtgaacacacccgctgtgcctcccaag 1424

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 181 TGGCCGGGTGGTCTCAAACTCTTGACCTCAAGTAACACCCGCTGTGCTCCCAAG 240

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

QY 1425 tgcgtgaattaccagcgtgagccaccatgccgggctcacacgctttgagttgataccattg 1484

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Aurora Biosciences Corp.
11010 Torreyana Road, San Diego, CA 92121, USA
Tel: 8584048436
Fax: 8584046719
Email: henkelg@aurorabio.com

Pools of cells were isolated from a Genomescreen(TM) library. The library of cells was generated by retroviral integration of a gene tagging element consisting of: 1) A promoterless beta-lactamase preceded by a splice acceptor as a reporter for gene expression; 2) A promoter driving neomycin resistance followed by a splice donor to trap downstream exons. 3' RACE from neomycin gene was performed using total RNA from isolated pools. Output was shotgun cloned in pAmp-1 and used to transform DH5-alpha competent bacteria. 5' ends of reported sequences were immediately preceded by splice donor from the trapping construct.

Class: exon-trapped.

FEATURES Location/Qualifiers

source 1..345
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Genetrap T47D Human Breast Carcinoma Library"
/tissue_type="Carcinoma"
/cell_type="Epithelial"
/cell_line="T47D"
/note="Organ: Breast; Vector: pAmp-1; 3' RACE of total RNA from genetrap pools; shotgun clone in pAmp-1 and used to transform DH5-alpha competent bacteria."
BASE COUNT 113 a 77 c 82 g 73 t

Query Match 14.9%; Score 301; DB 12; Length 345;
Best Local Similarity 98.9%; Pred. No. 2.9e-35;
Matches 345; Conservative 0; Mismatches 0; Indels 4; Gaps 4;

Qy 210 ccatactccattgcctccagctgttcacgcagggtttccacatcatatttccagaagcc 269
Db 1 CCATACCTTCCCATTTGCCCTCCAGCTGTTCCAGGAGTTTCACATCATATTTCCAGAAGGC 60
Qy 270 tcttgaagaagtagtaattgtcgcattccagagagctgatgggattgtgacttgctg 329
Db 61 TCTTGGAAA-AGTGATATATGTGTCATCCAGAGAGCTGATGGGATTTGTGACTTGGCTG 119
Qy 330 ctgtcatcttcattgcctccagcagaagaattctgtgtcagccgcacacacatactgtta 389
Db 120 CTGTCATCTTTCATGTCAA-CGCAGAAATCTGTGTGAGGAGCTGATGGGATTTGTGACTTGGCTG 178
Qy 390 agcagtgatgaagtgcaactgcctccagcagaagaattgttaagggaattgttgcacagga 449
Db 179 AGCAGTGGATGAAAGTGCAGCTGCCAAG-AAAATGTTAAAGGAAATGTTGCCACAGGA 237
Qy 450 agaaacaccatggcgaagggaacagtagtaacagggcacatcaggggaacacacgaaacatacg 509
Db 238 AGAACAACCATGCGCAGAGGAGACAGTACAGGGGCACATC-GGGGAACACCAACAAACATACG 296
Qy 510 gccataaaactcccttattagagagctcacagataaaatctacagagacaa 558
Db 297 GCCATAAAACTCCTTATTATAGAGTCTACAGATAAATCTACAGAGACAA 345

RESULT 15
BE004848/c
LOCUS BE004848 510 bp mRNA linear EST 05-JUN-2000
DEFINITION MR2-BN0114-020500-013-b12 BN0114 Homo sapiens cDNA, mRNA sequence.

ACCESSION BE004848
VERSION BE004848.1 GI:8265081
KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 510)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,W.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL MEDLINE

COMMENT

Contact: Simpson A.J.G.

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=6t2-MR2-BN0114-020

500-013-b12&t3=2000-05-02&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 21

High quality sequence stop: 442.

FEATURES

source

1..510

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="BN0114"

/dev_stage="Adult"

/note="Organ: breast_normal; Vector: puc18; Site:1: SmaI; Site:2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No.196,716 - Ludwig Institute for Cancer Research)

profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 176 a 109 c 104 g 121 t

ORIGIN

Query Match 14.5%; Score 292.6; DB 9; Length 510;
Best Local Similarity 94.1%; Pred. No. 4.3e-34;
Matches 304; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 671 ggtatgcacatgtagcacaataatactcaaacctcctggctcaagcgatccctccacct 730
Db 340 GGTCCTACTATGTGGCCAGGTGATCTCAAACTCCTGGGCTCGAGCGATCCTCCACCT 281
Qy 731 tagcctcccaaaagtactgggattataggtgtgagccacagtcgctgcctaattattttc 790
Db 280 TAGCCTCCCAAGTACTGGGATTTAGTGTGAGCCACAGTGCCTGGCCTAAATTTTTC 221
Qy 791 ttgtgatacaattcaggttttaattgttttggtaagaatttcctacgtgaattcgtgac 850
Db 220 TTGTGATCATATTCAGGTTTAAATGTTTGTAAAGAAATTTCTACGTGATTCGTGTAC 161
Qy 851 ttatttgcatttagagttcataaataattagggtttattttctataaataagattgtaa 910
Db 160 TTATTTTGCATTTAGAGTTTCATAAATATTAGGGTTTATTTCATAAATAGAAATTTAA 101
Qy 911 actaaataataactcacaacgcttagttttagtagtcacgctgtgttggattggaatttt 970
Db 100 ACTAAATATTAACCTCTAAACGCTAGTTTTCAGTAGTACCGCTGTTGGATTGAAATTTT 41
Qy 971 ctgatactgaaaagaacaaaag 993
Db 40 CTGATACTGAAAGAACATATAG 18

Search completed: October 7, 2002, 09:31:08
Job time: 1753 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 7, 2002, 09:01:55 ; Search time 58.07 Seconds
(without alignments) 8531.820 Million cell upd

Title: US-09-813-492-1

Perfect score: 2017

Sequence: 1 tagataccctgaacacctcc.....ataccaaaaaaaaaaaaaa 2017

Scoring table: IDENTITY NUC

scoring table: IDENTIFY_NOC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Minimum DB seq length:	0
Maximum DB seq length:	2000000000

Post-processing: Minimum Match 0%

Post-processing: Minimum Match 0%
Maximum Match 100%

Maximum match 100%
Listing first 45 summaries

Database :

```

Database :
Issued_Patents_NA:**
1: /cgn2.6/ptodata/1/ina/5A_COMB.seq:**
2: /cgn2.6/ptodata/1/ina/5B_COMB.seq:**
3: /cgn2.6/ptodata/1/ina/6A_COMB.seq:**
4: /cgn2.6/ptodata/1/ina/6B_COMB.seq:**
5: /cgn2.6/ptodata/1/ina/PCtUS_COMB.seq:**
6: /cgn2.6/ptodata/1/ina/backfiles1.seq:**

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	%			
1	1035.2	51.3	3117	4	US-09-146-580-6	Sequence 6, Appli
2	379	18.8	381	4	US-09-146-580-7	Sequence 7, Appli
3	302	15.0	311	4	US-09-146-580-11	Sequence 11, Appl
4	230	11.4	35060	3	US-08-814-095-7	Sequence 7, Appli
5	218.2	10.8	2839	4	US-09-061-702-1	Sequence 1, Appli
6	217	10.8	8453	4	US-09-167-681-45	Sequence 45, Appl
7	216.4	10.7	631	4	US-09-385-982-354	Sequence 354, Appl
8	213.2	10.6	1701	4	US-09-078-294-9	Sequence 9, Appli
9	212.8	10.6	87350	3	US-08-781-891-79	Sequence 79, Appl
10	212.2	10.5	841	5	PCF-US93-06251-80	Sequence 80, Appl
11	212.2	10.5	841	5	PCF-US93-06251-81	Sequence 81, Appl
12	212.2	10.5	22481	4	US-08-367-81A-43	Sequence 43, Appl
13	212.2	10.5	22481	5	PCF-US95-07201-43	Sequence 43, Appl
14	211.6	10.5	282	1	US-08-133-629-8	Sequence 8, Appli
15	211.2	10.5	17327	1	US-07-906-871-15	Sequence 15, Appl
16	211	10.5	8453	4	US-09-167-681-45	Sequence 45, Appl
17	210.8	10.5	10380	4	US-09-077-3548-3	Sequence 3, Appli
18	210.6	10.4	15977	4	US-09-608-285A-59	Sequence 59, Appl
19	210.4	10.4	6769	1	US-08-480-784-20	Sequence 20, Appl
20	210.4	10.4	6769	1	US-08-483-553-20	Sequence 20, Appl
21	210.4	10.4	6769	1	US-08-487-002-20	Sequence 20, Appl
22	210.4	10.4	6769	1	US-08-483-5548-20	Sequence 20, Appl
23	210.4	10.4	6769	1	US-08-488-011B-20	Sequence 20, Appl
24	210.4	10.4	6769	4	US-08-850-727-20	Sequence 20, Appl
25	210.4	10.4	6769	5	PCF-US95-10202-20	Sequence 20, Appl
26	210.4	10.4	6769	5	PCF-US95-10203-20	Sequence 20, Appl
27	210.4	10.4	6769	5	PCF-US95-10220-20	Sequence 20, Appl

28	210	10.4	53526	3	US-08-658-136-2	Sequence 2, Appli
29	210	10.4	53577	3	US-08-658-136-1	Sequence 1, Appli
30	209.4	10.4	36741	4	US-09-301-665-3	Sequence 3, Appli
31	209	10.4	14747	4	US-09-608-2875A-42	Sequence 42, Appli
32	207.8	10.3	20303	1	US-08-370-975B-6	Sequence 6, Appli
33	207.8	10.3	26764	1	US-08-370-975B-1	Sequence 1, Appli
34	207.6	10.3	2115	1	US-08-395-800A-7	Sequence 7, Appli
35	207.6	10.3	59065	4	US-09-813-817-3	Sequence 3, Appli
36	207.4	10.3	31571	1	US-08-323-443B-1	Sequence 1, Appli
37	207.2	10.3	4823	2	US-08-457-254-5	Sequence 5, Appli
38	207.2	10.3	4823	2	US-08-484-257-20	Sequence 20, Appli
39	207.2	10.3	4823	3	US-08-999-927-5	Sequence 5, Appli
40	207.2	10.3	4823	4	US-08-461-818-5	Sequence 5, Appli
41	207.2	10.3	4823	5	PCR-US94-08806-28	Sequence 28, Appli
42	207.2	10.3	4823	5	PCR-US95-01829-5	Sequence 5, Appli
43	207.2	10.3	4823	5	PCR-US95-16626-5	Sequence 5, Appli
44	206.8	10.3	13865	3	US-09-009-217-11	Sequence 11, Appli
45	206.8	10.3	13865	3	US-09-009-656-11	Sequence 11, Appli

ALIGNMENTS

RESULT 1

RESUL I
US-09-146-580-6
: Sequence 6, Application US/09146580A

; Patent No. 6306653

; PATIENT NO. 000000000
; GENERAL INFORMATION:

APPLICANT: Papsidero, Lawrence

APPLICANT: Dyster, Lyn M

APPLICANT: Frustaci, Jana M

; TITLE OF INVENTION: DETECTION

FILE REFERENCE: 200755/1002

; CURRENT APPLICATION NUMBER: 1888-8

; CURRENT FILING DATE: 1998-0

; EARLIER APPLICATION NUMBER:

; EARLIER FILING DATE: 1998-0...

; EARLIER APPLICATION NUMBER:

EARLIER FILING DATE: 1998-01-19
NUMBER OF SEC ID NOS: 19

```

; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Datascan Ver. 3.0

```

```

; SOFTWARE: PatentIn Ver. 2.0
; CEO TD NO 6

```

```

; SEQ ID NO 6
:   LENGTH: 3117

```

```

; LENGTH: 311/
; TYPE: DNA

```

```

; TYPE: DNA
; ORGANISM: Homo sapiens

```

```

; ORGANISM: HOMO SAPIENS
: FEATURE:

```

```

; FEATURE:
; NAME/KEY: unsure

```

```

; NAME/REL: unsure
; LOCATION: (35)

```



```
; Patent No. 6306653
; GENERAL INFORMATION:
; APPLICANT: Papsidero, Lawrence D
; APPLICANT: Dyster, Lyn M
; APPLICANT: Frustaci, Jana M
; TITLE OF INVENTION: DETECTION AND TREATMENT OF BREAST DISEASE
; FILE REFERENCE: 200755/1002
; CURRENT APPLICATION NUMBER: US/09/146,580A
; CURRENT FILING DATE: 1998-09-03
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 60/071,889
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 60/092,155
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 311
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (101)
; OTHER INFORMATION: N at position 101 is either A, C, G, or T
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (162)
; OTHER INFORMATION: N at position 162 is either A, C, G, or T
; US-09-146-580-11

Query Match      15.0%; Score 302; DB 4; Length 311;
Best Local Similarity 99.3%; Pred. No. 6.5e-59;
Matches 302; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 208 agccatactccattgctccagctgttgccagaggttccacatcatattccagaag 267
Db 311 ACCCATACTTCCCATGCTCCAGCTGTTCACGGAGGTTTCACATCATATTTCCAGAAG 252
QY 268 gctctgaaagagtgaaatgtctgcacatccagagctgatgggattgacttgc 327
Db 251 GCTCTGGAAAGAGTGAATATGTGCGCATCCAGAGCTGATGGGATTGTGACTTGGC 192
QY 328 tgctgtcatcttctatgtcaagcgcagaaagaatctgtgtagccgcgcacacatactgt 387
Db 191 TGCTGTGCATCTTCATGTCAAGCGCAGGAANAATCTGTGTCAGCCGCCACACCATACTCT 132
QY 388 taagcagtgagtgaaagtgagctgcagagcgaagaaatgtaagaaatgtttgccacag 447
Db 131 TAAGCAGTGGATGAAGTGCAAGCTGCCAANAATAATGTAAGGAAATGTTTGCACAG 72
QY 448 gaagaaacaccatggcagagagaacagtaacagggcacatcagggaggaacacagaacata 507
Db 71 GAAGAAACACCATGGCAAGAGGAAACAGTAACAGGGCCATCAGGGGAAACACGAACATA 12
QY 508 cggc 511
Db 11 CGGC 8

RESULT 4
US-08-814-095-7/c
; Sequence 7, Application US/08814095
; Patent No. 6025183
; GENERAL INFORMATION:
; APPLICANT: Soreq, Hermona
; APPLICANT: Zakut, Haim
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR
; TITLE OF INVENTION: ANTI-CHOLINESTERASE SUBSTANCES
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KOHN & ASSOCIATES
; STREET: 30500 No. 6025183thwestern Highway, Suite 410

; CITY: Farmington Hills
; STATE: Michigan
; COUNTRY: U.S.
; ZIP: 48334
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/814,095
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Montgomery, Ilene N.
; REGISTRATION NUMBER: 38,972
; REFERENCE/DOCKET NUMBER: 2391.00066
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (248) 539-5050
; TELEFAX: (248) 539-5055
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35060 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Cosmid including ACHE"
; DESCRIPTION: promotor, ACHE gene and ARS gene"
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: 7q22
; FEATURE:
; NAME/KEY: promotor
; LOCATION: 4089..22464
; OTHER INFORMATION: /function= "ACHE Promotor"
; OTHER INFORMATION: /standard_name= "ACHE Promotor"
; FEATURE:
; NAME/KEY: exon
; LOCATION: 22465..22537
; OTHER INFORMATION: /function= "non-translated"
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 1
; FEATURE:
; NAME/KEY: exon
; LOCATION: 24090..25177
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /function= "(translation start:"
; OTHER INFORMATION: 24110)"
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 2
; FEATURE:
; NAME/KEY: exon
; LOCATION: 25524..26009
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 3
; FEATURE:
; NAME/KEY: exon
; LOCATION: 27005..27274
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /gene= "ACHE"
; OTHER INFORMATION: /number= 4
; FEATURE:
; NAME/KEY: exon
; LOCATION: 27255..28007
; IDENTIFICATION METHOD: experimental
```



```

; NAME: Greer, Helen
; REGISTRATION NUMBER: 36,816
; REFERENCE/DOCKET NUMBER: M0828/7001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; TELEX: 92-1742 EZEKEL
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 282 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-133-629-8

```

Query Match	10.5%	Score 211.6;	DB 1;	Length 282;
Best Local Similarity	85.1%;	Pred. No. 9.6e-39;		
Matches 240;	Conservative 6;	Mismatches 35;	Indels 1;	Gaps 1;

QY	1178	gagacagagtctcaactctgtttgccaggctggagtcgaagtgcacaatcctcggtccattg	1237
		::: ::: ::: ::: ::: ::: ::: :::	
Db	282	GAGAYRAGTCTYRCTCTGTTGCCAGGCTGGAGTGCGAGTCWGGGTCACTG	223
QY	1238	caacctccgctcccgggtccaagtattctcttgcctcagccctccccaaagtaactgat	1297
		::: ::: ::: ::: ::: ::: :::	
Db	222	CARCTCCACCTCTCGGGTTC AAGCGATTCCTCGCCTCACGCCCTCCGAGTAGTGGGAT	163
QY	1298	tacaggcgccagaccacaccccgcgtgattttgtatttttagtagagacgggggtttt	1357
		::: ::: ::: ::: ::: ::: :::	
Db	162	TACAGGCGGC -GCCACCACGCCCGCGTAATTTTTGTATTTTTATTAGTAGACAGGGGTTTC	104
QY	1358	cccacgttgccggctgggtctcaaacctcttgaocccaagtgaaccacccgcctgtgcct	1417
		::: ::: ::: ::: ::: ::: :::	
Db	103	ACCATGTTGGCCAGGCTGGTCTTGAACCTCCTCGACCTCAGGTGATCCACCCACCTCGGCCT	44
QY	1418	cccaaatgctgtaattacacagcgtgagccaccatgcggggc	1459
		::: ::: ::: ::: ::: ::: :::	
Db	43	CCCAAAGTGTGGGATTATACAGTGTGAGCCACCAACGCGCCAGC	2

RESULT 15
 US-07-906-871-15/c
 ; Sequence 15, Application US/07906871
 ; Patent No. 5340739
 ; GENERAL INFORMATION:
 ; APPLICANT: Stevens, Richard L.
 ; APPLICANT: Avraham, Shalom
 ; TITLE OF INVENTION: HEMATOPOIETIC CELL SPECIFIC
 ; TITLE OF INVENTION: TRANSCRIPTIONAL REGULATORY ELEMENTS OF SERGLYCIN AND USES
 ; TITLE OF INVENTION: THEREOF
 ; NUMBER OF SEQUENCES: 18
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
 ; STREET: 1225 Connecticut Avenue, N.W., Suite 300
 ; CITY: Washington
 ; STATE: DC
 ; COUNTRY: USA
 ; ZIP: 20036
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/906,871
 ; FILING DATE: 19920103
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/816,289
 ; FILING DATE: 03 JAN 1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/635,544

```

1 FILING DATE: 18-JAN-1991
2
3 PRIOR APPLICATION DATA:
4 APPLICATION NUMBER: PCT/US89/03051
5
6 FILING DATE: 13-JUL-1989
7
8 PRIOR APPLICATION DATA:
9 APPLICATION NUMBER: US/07/224,035
10
11 FILING DATE: 13-JUL-1988
12
13 ATTORNEY/AGENT INFORMATION:
14 NAME: Cimbala, Michele A
15
16 REGISTRATION NUMBER: 33,851
17
18 REFERENCE/DOCKET NUMBER: 0627.2830004
19
20 TELECOMMUNICATION INFORMATION:
21 TELEPHONE: (202)833-7533
22
23 TELEFAX: (202)833-8716
24
25 INFORMATION FOR SEQ ID NO: 15:
26
27 SEQUENCE CHARACTERISTICS:
28 LENGTH: 17327 base pairs
29
30 TYPE: NUCLEIC ACID
31
32 STRANDEDNESS: both
33
34 TOPOLOGY: linear
35
36 MOLECULE TYPE: DNA
37
38 FEATURE:
39
40 NAME/KEY: exon
41
42 LOCATION: 621..753
43
44 FEATURE:
45
46 NAME/KEY: intron
47
48 LOCATION: 754..9596
49
50 FEATURE:
51
52 NAME/KEY: exon
53
54 LOCATION: 9597..9744
55
56 FEATURE:
57
58 NAME/KEY: intron
59
60 LOCATION: 9745..16396
61
62 FEATURE:
63
64 NAME/KEY: exon
65
66 LOCATION: 16397..17327
67
68 US-07-906-871-15

```

	Query Match	10.5%	Score 211.2;	DB 1;	Length 17327;
	Best Local Similarity	82.5%;	Pred. No. 3,7e-38;		
Matches	254;	Conservative	0;	Mismatches 53;	Indels 1; Gaps 1;
QY	1157	gtgcttgggtgttttttttgagacagagtctcaactctattgccaggctggagtcag	1216		
Db	7833	GTACTTTTTTTTTTTTGTAGATGAAGTCTCGCTCTGTTGCCCTGGCTGGGATGCAG	7774		
QY	1217	tggcacaatctcggtcatgtgaacctccgcctccgcgttcaaagtattctcttgcctc	1276		
Db	7773	TGGCACAAATTCGGCTCACA'TGCACACTCTGCCCTCCAGGTTCAAGCCATTCTCTGGCTC	7714		
QY	1277	agctcccagaagtaactgatattacaggcgccccagccaccacccccgcgtgatttttgtat	1336		
Db	7713	AGCCTCCCGAGTAGCTGGGATGACAGGTGCCCA -CAATCATATCCAGTAATTTTGTAT	7655		
QY	1337	tttttagtagacagcgggtttttccacgttggccggggtggtctcaaaactcttgacctcaa	1396		
Db	7654	TTTTAGTAGGATGGGGTTTTCACCATTGTGCCAGGCTGGTCTCGAACTCTTGACCTCAG	7595		
QY	1397	gtgaaccaccgcgcgtgcctcccaaatgctgtggaattaccagcgtgagccaccatgcc	1456		
Db	7594	GTGATCCACCCGGCTCAGTCTCCCAAAGTCTGGGATTACAGGTGTGAGCCACCACACC	7555		
QY	1457	ggctcac	1464		
Db	7534	GGCGAAA	7527		

Search completed: October 7, 2002, 09:36:02
Job time: 2047 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 7, 2002, 09:01:55 ; Search time 2534.54 Seconds
(without alignments)
16653.446 Million cell updates/sec

Title: US-09-813-492-1
Perfect score: 2017
Sequence: 1 tagataccctgaacacctcc.....ataccaaaaa..... 2017

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg_hum.*
31: em.htg_inv.*
32: em.htg_other.*
33: em.htgo_inv.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	Score	Match	Length	DB	ID	Description
No.	Match	Length	DB	ID			

1	1314	65.1	107819	9	AC010465	AC010465 Homo sapi
2	1314	65.1	140734	9	AC025457	AC025457 Homo sapi
3	1314	65.1	143372	2	AC022132	AC022132 Homo sapi
4	1035.2	51.3	3117	6	ARI74328	ARI74328 Sequence
5	497	24.6	768	9	AF266504	AF266504 Homo sapi
6	497	24.6	1349	9	AF110384	AF110384 Homo sapi
7	384	19.0	384	9	AF220210	AF220210 Homo sapi
8	379	18.0	381	6	ARI74329	ARI74329 Sequence
9	361.8	17.9	472	6	AX329666	AX329666 Sequence
10	361.8	17.9	472	6	AX334850	AX334850 Sequence
11	302	15.0	311	6	ARI74333	ARI74333 Sequence
12	243.6	12.1	61101	2	AC083975	AC083975 Homo sapi
13	243.6	12.1	146376	9	AC009247	AC009247 Homo sapi
14	242	12.0	22426	2	AL105916	AL105916 Homo sapi
15	242	12.0	149916	2	AL359540	AL359540 Homo sapi
16	241.6	12.0	53402	9	AL158844	AL158844 Human DNA
17	241.6	12.0	112219	2	AL390315	AL390315 Homo sapi
18	241.4	12.0	100998	9	HS343K2	AL031659 Human DNA
19	240.2	11.9	146443	2	AC067726	AC067726 Homo sapi
20	240.2	11.9	149899	2	AC074266	AC074266 Homo sapi
21	240.2	11.9	190069	9	AC011597	AC011597 Homo sapi
22	240	11.9	141289	9	AL449363	AL449363 Human DNA
23	239.6	11.9	174231	9	AC021054	AC021054 Homo sapi
24	239.6	11.9	213464	9	AC022165	AC022165 Homo sapi
25	238.8	11.8	135044	9	AC006001	AC006001 Homo sapi
26	238.6	11.8	41369	9	CH19F15314	AD000091 Homo sapi
27	238.2	11.8	149425	2	AC013320	AC013320 Homo sapi
28	238.2	11.8	191754	9	AC021016	AC021016 Homo sapi
29	238	11.8	86428	9	AL390776	AL390776 Human DNA
30	238	11.8	93821	9	AC004254	AC004254 Homo sapi
31	238	11.8	153937	2	AC023953	AC023953 Homo sapi
32	237.8	11.8	37680	2	AL359985	AL359985 Homo sapi
33	237.4	11.8	35016	9	AL603838	AL603838 Human DNA
34	237.4	11.8	133273	2	AC068586	AC068586 Homo sapi
35	237.4	11.8	144555	2	AC037451	AC037451 Homo sapi
36	237.4	11.8	165394	2	AF000877	AF000877 Homo sapi
37	237.4	11.8	168729	2	AL590137	AL590137 Homo sapi
38	237.4	11.8	176405	9	AL138768	AL138768 Human DNA
39	237.4	11.8	178601	9	HSJ1022P6	AL109835 Human DNA
40	237.2	11.8	87424	9	AC006077	AC006077 Homo sapi
41	237.2	11.8	88698	2	AC009008	AC009008 Homo sapi
42	237.2	11.8	174526	9	AF001183	AF001183 Homo sapi
43	237.2	11.8	180964	2	AC008955	AC008955 Homo sapi
44	237.2	11.8	190314	9	AC019227	AC019227 Homo sapi
45	237	11.8	214287	2	AC040169	AC040169 Homo sapi

ALIGNMENTS

RESULT	1	AC010465	107819 bp	DNA	linear	PRI 08-JUN-2001
AC010465	LOCUS	Homo sapiens chromosome 5 clone CTD-2282F8, complete sequence.				
DEFINITION	AC010465					
ACCESSION	AC010465.7	GI:14329074				
VERSION	HTG.					
KEYWORDS	human.					
SOURCE	Homo sapiens					
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1 (bases 1 to 107819)					
AUTHORS	DOE Joint Genome Institute and Stanford Human Genome Center.					
TITLE	Direct Submission					
JOURNAL	Unpublished					
REFERENCE	2 (bases 1 to 107819)					
AUTHORS	DOE Joint Genome Institute.					
TITLE	Direct Submission					
JOURNAL	Submitted (15-SEP-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA					
REFERENCE	3 (bases 1 to 107819)					
AUTHORS	DOE Joint Genome Institute and Stanford Human Genome Center.					
TITLE	Direct Submission					


```
REFERENCE
AUTHORS   4 (bases 1 to 140734)
TITLE     DOE Joint Genome Institute and Stanford Human Genome Center.
JOURNAL   Direct Submission
COMMENT   Submitted (08-JUN-2001) DOE Joint Genome Institute, 2800 Mitchell
         Drive, Walnut Creek, CA 94598, USA
         On Jun 8, 2001 this sequence version replaced gi:13470150.
         Draft Sequence Produced by DOE Joint Genome Institute
         www.jgi.doe.gov
         Finishing Completed at Stanford Human Genome Center
         www.shgc.stanford.edu
         Quality: Phrap Quality >=40 99.5% of Sequence;
         Estimated Total Number of Errors is 0.3.
         STS Content:
         SHGC-64113 G38695.

FEATURES             Location/Qualifiers
     source            1..140734
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="CTD-2202K16"
BASE COUNT  43264 a 28409 c 28508 g 40553 t
ORIGIN

Query Match      65.1%; Score 1314; DB 9; Length 140734;
Best Local Similarity 98.9%; Pred. No. 3e-292;
Matches 1323; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY  671 ggtatgcaaatgagccaaataataactcaaaactctgggtcaagcgatctctccacact 730
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB  31533 GGTCTCACTATGTCGCCAGGTGTATCTCAAACTCTCTGGCTCAAGCGATCTCCACCT 31474

QY  731 tagctcccaagtagctggattataggtgtagcacaagtagtcctggtcctaatttttc 790
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB  31473 TAGCTCCCAAGTAGTCTGGGATATAGGTGTGAGCCACAGTGCCTGGCCCTAATATTTC 31414

QY  791 ttgtgatcaaatcaggtttaaagtgttttggtaagaatttctacgtgaattcgtgtac 850
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB  31413 TTGTGATCAAAATCAGGTTAATGTTTGGTTAGAAATTTCTAGTGAATTCGGTATC 31354

QY  851 ttattttgtcatttagagttcataaattatagggtttatttcttaataagaatgtttaa 910
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB  31353 TTATTTTGTCATTAGATTGATATAAATATTAGGGTTTATTTCCTAAATAGAAATGTTAA 31294

QY  911 actaaataataactcaaaacgctagtgttagtagtagtaaccgtgtgttgattgaaatttt 970
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB  31293 ACTAAATATAACTTCAAAACGCTAGTTGAGTAGTACCGTGTGTTGGATTGAAATTTT 31234

QY  971 ctgatactgaaagaaacaaagcctgccttctgcccagaaccttttgcctccccagct 1030
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB  31233 CTGATACTGAAAGAACAAAAGCCTGCTTTCTGCCAGAACCTTTTGCCTCCCCCAGT 31174

QY  1031 cagttcttggagcagcactagttaggggcccagagttcgcccttctgtgtgatttta 1090
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB  31173 CAGTTCTTGGAGCAGCAGTATTAGGGGCCAGAGTTCGGCCTCTGTGTGGTGAATTTA 31114

QY  1091 cgtctgcttaacaagaagcctacatctttagctcctattccaccttctcaacgctt 1150
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB  31113 CGCTCTGCCCTAAACAGGAGCCTACATCTTTTAGTCTCTTATTCACCCCTTCTCACACGTT 31054

QY  1151 ttgttgttgttgggtgtgttttttttttttttttttttttttttttttttttttttttt 1210
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB  31053 TTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 30994

QY  1211 gtgcagtggcacaatctcgggtcatttgaacctccgcctcccgcttcaagtattctct 1270
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB  30993 GTGCAGTGGCACAATCTCGGCTCATTGTCAACCTCCGCCCTCCCGGTTCAAGTGATTCTCT 30934

QY  1271 tgcctcagcctcccaagtaactgattatcacaggcgcacacacacaccccgctgattt 1330
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB  30933 TGCCCTCAGCCTCCCAAGTAACTGATATTACAGCGGCCAGCCACACACCCCGCTGATT 30874

QY  1331 ttgtatttttagagacgggggtttttccacgcttggcgggctggtctcacaactcttga 1390
```

```
Db  30873 TTGTATTTTCTAGTAGAGCGGGTTTTCACCGTGGCGGGCTGCTCTCAAACTCTTGA 30814
QY  1391 cctcaagtgaaacaccccgctgtgctcccaagtgctggaattaccagcgtgagccacc 1450
Db  30813 CCTCAAGTGAACACACCCGCTGTGCTCCCAAAGTGTCTGGAATTTACACAGCGTGAGCCACC 30754
QY  1451 atgccgggctcacagcttggattgataccattgtgccattctctcttggcctctttt 1510
Db  30753 ATGCCGGGCTCACACGTTTGAGTTGATACATTGTGCCATTCTCTTTGGGCTCTTTT 30694
QY  1511 tgtccatagaggtctcaagatagataggtgaagagccagtagtggttcataaagaagccaat 1570
Db  30693 TGTCCATAGAGGCTTCAAGATAGATAGTAAAGAGCCAGTAGTGTTCATAAAGAAGCAAT 30634
QY  1571 agagagcaggagccactttatcaggtggcaggtgtcccggggctccctcctgctgagctacc 1630
Db  30633 AGAGAGCAGGAGGCCACTTTATCAGGTGGCAGGTGCTCTGGGCTCTCCCTGCTAGTACC 30574
QY  1631 caagcgtgtgtgtgccaggtctcttgaggtgataatgggacacacagagagcactgag 1690
Db  30573 CAAGCGGTGCTGTGCCAGGATGCTTGGAGGTGATAATGGGACACACAGAGGCACTGAG 30514
QY  1691 tctccataggttaaaatgccacaaaactggcctttgcctaatatccctcattgactatt 1750
Db  30513 TCTCCATAGGTTAAAAATGCCACAAAACCTGGCTTTGCCCTAATATATCCCTCATTTGACTATT 30454
QY  1751 tagcatttaatttatttcttcccgacatttctgcaagcttctgtaatttatatttccact 1810
Db  30453 TAGCATTTAAATTTATTTATTTTCTTGACATTTCTGCAAGCTTTGTATTTATTTTCCACT 30394
QY  1811 ttatagatagagaatttgagctcttagaggtcaaatgacttgcacaggtcacacagga 1870
Db  30393 TTATAGATGAGGAATTTTGAGGCTCTTAGAGGTAAATGACTTGCCACGCTCACAGGA 30334
QY  1871 agtggcagacagaagctttttaaagaagaaaaatttaataataataatagagagtaac 1930
Db  30333 AGTGGCAGAGACAAGCTTTTAAATAAGAAAAAATTAATAAATAATAATATGAGAGTAAC 30274
QY  1931 ttaaaataataaaacacaaattttaataataataacgctgataacccaacataataaa 1990
Db  30273 TTAATAATATAATAAACCAACAATTTTAAATTAATTAACCGTGATAACCAACATTAATAA 30214
QY  1991 agttaagatacacaacaaa 2008
Db  30213 AGTTAGATACCAAAACA 30196
```

RESULT 3

```
AC022132
LOCUS      Homo sapiens chromosome 5 clone CTD-220J22, WORKING DRAFT
DEFINITION
SEQUENCE, 9 unordered pieces.
ACCESSION  AC022132
VERSION    AC022132.5 GI:13699628
KEYWORDS   HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 143372)
AUTHORS     DOE Joint Genome Institute.
TITLE       Sequencing of Human Chromosome 5
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 143372)
AUTHORS     DOE Joint Genome Institute.
TITLE       Direct Submission
JOURNAL     Submitted (26-JAN-2000) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
            On Apr 20, 2001 this sequence version replaced gi:7711706.
            -----Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
```


[illegible]


```

AC083975
LOCUS       AC083975               61101 bp    DNA    linear    HTG 08-OCT-2000
DEFINITION Homo sapiens chromosome 8 clone RP11-391E1 map 8, LOW-PASS SEQUENCE
SAMPLING.
AC083975
VERSION     AC083975.1   GI:10717242
KEYWORDS    HTG; HTGS_PHASE0.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 61101)
AUTHORS     Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE       Homo sapiens chromosome 8, clone RP11-391E1
JOURNAL
REFERENCE   2 (bases 1 to 61101)
AUTHORS     Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bida,F., Boguslavsky,L.,
Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
Choepe,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,
FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,
Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L.,
Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., LaRocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G.,
Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K.,
McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,
O'Donnell,P., O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K.,
Pieroni,N., Pisanic,C., Pollara,V., Raymond,C., Rieback,M., Riley,R.,
Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,
Sougnuez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Tirell,A., Travers,M., Triglio,J., Vassiliev,H., Viel,R., Vo.A.,
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
Zimmer,A. and Zody,M.
JOURNAL     Direct Submission
COMMENT     Submitted (08-OCT-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: F11332
Center clone name: 391_E_1
-----
* NOTE: This record contains 76 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
1 720: contig of 720 bp in length
721 820: gap of 100 bp
821 1525: contig of 705 bp in length
1526 1625: gap of 100 bp
1626 2334: contig of 709 bp in length
2335 2434: gap of 100 bp
2435 3143: contig of 709 bp in length
3144 3243: gap of 100 bp
3244 3945: contig of 702 bp in length
3946 4045: gap of 100 bp
4046 4727: contig of 682 bp in length
4728 4827: gap of 100 bp
4828 5512: contig of 685 bp in length
5513 5612: gap of 100 bp
5613 6302: contig of 690 bp in length
6303 6402: gap of 100 bp
6403 7111: contig of 709 bp in length
7112 7211: gap of 100 bp
7212 7921: contig of 710 bp in length
7922 8021: gap of 100 bp
8022 8729: contig of 708 bp in length
8730 8829: gap of 100 bp
8830 9532: contig of 703 bp in length
9533 9632: gap of 100 bp
9633 10373: contig of 741 bp in length
10374 10473: gap of 100 bp
10474 11201: contig of 728 bp in length
11202 11301: gap of 100 bp
11302 11997: contig of 696 bp in length
11998 12097: gap of 100 bp
12098 12804: contig of 707 bp in length
12805 12904: gap of 100 bp
12905 13633: contig of 729 bp in length
13634 13733: gap of 100 bp
13734 14422: contig of 689 bp in length
14423 14522: gap of 100 bp
14523 15229: contig of 707 bp in length
15230 15329: gap of 100 bp
15330 16016: contig of 687 bp in length
16017 16116: gap of 100 bp
16117 16823: contig of 707 bp in length
16824 16923: gap of 100 bp
16924 17637: contig of 714 bp in length
17638 17737: gap of 100 bp
17738 18447: contig of 710 bp in length
18448 18547: gap of 100 bp
18548 19244: contig of 697 bp in length
19245 19344: gap of 100 bp
19345 20022: contig of 678 bp in length
20023 20122: gap of 100 bp
20123 20817: contig of 695 bp in length
20818 20917: gap of 100 bp
20918 21625: contig of 708 bp in length
21626 21725: gap of 100 bp
21726 22437: contig of 712 bp in length
22438 22537: gap of 100 bp
22538 23265: contig of 728 bp in length
23266 23365: gap of 100 bp
23366 24086: contig of 721 bp in length
24087 24186: gap of 100 bp
24187 24892: contig of 706 bp in length
24893 24992: gap of 100 bp
24993 25716: contig of 724 bp in length
25717 25816: gap of 100 bp
25817 26509: contig of 693 bp in length
26510 26609: gap of 100 bp
26610 27313: contig of 704 bp in length
27314 27413: gap of 100 bp
27414 28117: contig of 704 bp in length
28118 28217: gap of 100 bp
28218 28917: contig of 700 bp in length
28918 29017: gap of 100 bp
29018 29734: contig of 717 bp in length
29735 29834: gap of 100 bp
29835 30531: contig of 697 bp in length
30532 30631: gap of 100 bp
30632 31351: contig of 720 bp in length
31352 31451: gap of 100 bp
31452 32156: contig of 705 bp in length
32157 32256: gap of 100 bp
32257 32959: contig of 703 bp in length
32960 33059: gap of 100 bp
33060 33732: contig of 673 bp in length
33733 33832: gap of 100 bp

```

[illegible]

[illegible]

BASE COUNT 43232 a 34262 c 32167 g 40250 t 5 others
ORIGIN

/note="assembly_fragment:01220.0"

```
Query Match      12.0%; Score 242; DB 2; Length 149916;
Best Local Similarity 58.3%; Pred. NO. 2e-45;
Matches 462; Conservative 0; Mismatches 325; Indels 5; Gaps 2;

QY 668 tcggtatgcaaatagccaataataactcaataactcctgggctcaagcgatcctccca 727
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8602 TGGGATCTTGCCATGTTGCCAGGTGCTCAAACTCTGGGCTCAAGTGATCCTCTTG 8661
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 728 ccttagctcccaagtagctggtataggtgtagccacagtgctggcctaattatt 787
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8662 CCTCAGCTCCCAAGACTAGGATAGAGCATGAGCCACCATGCCAGCCAAAACATCC 8721
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 788 ttcttgatcaaaatcagggtttaattgttttgggtaagaatttctacgtgaattcgtg 847
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8722 TTTCTTACAAATCACCCAGTCTCAGGTATTTCTTTTCTTTCTTTCTTTTCTTTT 8781
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 848 taattatttgcattagaggttcataaattattagggtttattttctaaataagaattg 907
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8782 TTTTGTGAGAGAGGATTTGTCTCTCTACCTCAGGCTGGAGTGCAGTGGAAACAATCT 8841
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 908 taaactaaataataacttcaaaacgtctagttgtagtacctggtgtttggattgaaat 967
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8842 CAGCTCAGTGCACCTCCGCTCTCCGGTTCAGGAATCTTCCCATCTCAGGCTCCCCAG 8901
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 968 ttctgtactgaaagaacaaaagcctgccttcttgcgccagaacccctttgcctcccc 1027
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8902 CAGCTGAGACT----ACAGGCACACACCACCACTGGCTAATTTTATTTTCTGTGA 8957
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1028 agtcagttcttgagcagcactagttaggggcccagagattcgccctctctgtgtgatt 1087
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8958 ATTTTGTGAGAGATGAGGTTTCTCCATGTTGCCCAGGCTGGAGGAATTTCTAGCAGTGC 9017
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1088 ttacgctctgcctaaacaggagcctacatcttttagctcctattccaccctctcacac 1147
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9018 AAGAGCACCTTAATACATTCCTTTAAATCAGAGAAAAAACAACCTTTTAAGTCARAA 9077
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1148 gttttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1207
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9078 GTTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 9137
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1208 ggagtgagtgccacaatctcggtcatttgcaacctccgctcccgcttcaagtgttc 1267
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9138 GGAGTGCAATGCATGATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 9197
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1268 tcttgctcagctcccaagtaactgatattacagcgccagccacacaccccgctga 1327
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9198 TCCTGCCTGAGCCTCCCAAGTAGCTGGAATATACAGGTGCACA-CCACCACCCAGCTAA 9256
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1328 ttttgtatttttagtagagcgggttttccacagttggccgggtgtgtctcaaacctct 1387
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9257 TTTTGTGATTTTGTAGAGACGGGTTTTCACCATGTTCCCCAGACTGCTCTCGAACATCC 9316
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1388 tgacctcaagtgaacaccccgctgtgctcccaagtgtgtgaattaccagcgtgagcc 1447
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9317 TGACCTCAAGTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 9376
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1448 accatgcggggc 1459
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9377 ACTGCACCCAGC 9388
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: October 7, 2002, 10:28:27
Job time: 5192 sec

QY 512 cataaaactccttattagagtgctacagataaattacagagacaattcctcaagtga 571
|||||
Db 421 CATAAAACCTCTTATTAGAGTCTACAGATAAATCTACAGACAATTCTCAAGTGA 480
QY 572 ctiggccatgattggtt 588
|||||
Db 481 CTTGGCCCATGATTGGTT 497
RESULT 6
AF220210 AF220210 384 bp mRNA linear PRI 23-JUL-2000
LOCUS Homo sapiens CC chemokine CCL28 (SCYA28) mRNA, complete cds.
AF220210
VERSION AF220210.1 GI:9392590
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 384)
AUTHORS Wang, W., Soto, H., Oldham, E.R., Buchanan, M.E., Homey, B., Catron, D.,
Jenkins, N., Copeland, N.G., Gilbert, D.J., Nguyen, N., Abrams, J.,
Kershenovich, D., Smith, K., McClanahan, T., Vicari, A.P. and
Zlotnik, A.
TITLE Identification of a novel chemokine (CCL28), which binds CCR10
JOURNAL J. Biol. Chem. 275 (29), 22313-22323 (2000)
MEDLINE 20357357
REFERENCE 2 (bases 1 to 384)
AUTHORS Soto, H., Vicari, A. and Zlotnik, A.
TITLE Direct Submission
JOURNAL Submitted (30-DEC-1999) Immunobiology, DNAX Research Institute, 901
California Ave, Palo Alto, CA 94304, USA
FEATURES
source 1. .384
/organism="Homo sapiens"
/db_xref="taxon:9606"
gene 1. .384
/gene="SCYA28"
CDS 1. .384
/gene="SCYA28"
/codon_start=1
/product="CC chemokine CCL28"
/protein_id="AAF87205.1"
/db_xref="GI:9392591"
/translation="MQORGLAIVLAVCAALHASRAILPIASCTEYSHLSRLLE
RYNMRIQADGCDLAIVLHVKKRRICVSPHNITVQWMKVQAARKNGKNGVCHRK
KHHGKRNSRAHQGHETHYGHKTPY"
BASE COUNT 113 a 91 c 99 g 81 t
ORIGIN
Query Match 19.0%; Score 384; DB 9; Length 384;
Best Local Similarity 100.0%; Pred. No. 3.5e-199;
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 146 atgcagcagagagactcgccatggtggccttgctgtctgtcgccctacatgcctca 205
Db 1 ATGCAGCAGAGAGACTCGCCATCGTGGCCCTTGCGCTGTCTGTGCGGCCCTACATGCTCA 60
QY 206 gaagccatacttcccattgctccagctgtgacgagaggtttcacatcatattccaga 265
Db 61 GAAGCCATACTTCCCATTCCTCCAGCTGTTCACGCGAGGTTTCACATCATATATTCAGA 120
QY 266 aggtcctcgaagagatgaatatgtctgcataccagagagctgatgggattgtgacttg 325
Db 121 AGGCTCTGGAAGAGTGAATATGTCTGCATCCAGAGAGCTGATGGGATTGTGACTTG 180
QY 326 gctgctgctatccttcattgctcaagccagagaagatctgtgtcagcccgacacaccatact 365
Db 181 GCTGCTGCTATCTTCTATGTCACGCCAGAGAATCTGTGTGTCAGCCGCGACCACTACT 240
QY 386 gtttaagcagtgatgaagtgcagctgccaaagaaaaatggttaagaaaatgttggccac 445

Db 241 GTTAACAGTGGTGAAGTGCAGGCTGCCAGAAAAATGGTAAAGAAATGTTTCCAC 300
QY 446 agaaacaaacaccatgccaagaggaacagtaaacagggcacatcaggggaaacacgaaca 505
|||||
Db 301 AGGAAGAAACACCATGGCAAGAGGAACAGTAACAGGGCACATCAGGGGAAACACGAACA 360
QY 506 tacggccataaaaaactccttattag 529
|||||
Db 361 TAGGGCCATAAAACTCTTTATTAG 384
RESULT 7
AR174328 AR174328 3117 bp DNA linear PAT 17-DEC-2001
LOCUS Sequence 6 from patent US 6306653.
DEFINITION AR174328
ACCESSION AR174328
VERSION AR174328.1 GI:17914648
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3117)
AUTHORS Papsidero, L.D., Dyster, L.M. and Frustaci, J.M.
TITLE Detection and treatment of breast disease
JOURNAL Patent: US 6306653-A 6 23-OCT-2001;
FEATURES Location/Qualifiers
source 1. .3117
/organism="unknown"
BASE COUNT 860 a 582 c 566 g 921 t 188 others
ORIGIN
Query Match 15.5%; Score 313; DB 6; Length 3117;
Best Local Similarity 100.0%; Pred. No. 3.3e-160;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 697 ctcaaaactcctgggtcgaagcgatcctccacaccttagcctcccaagtagtactgggattata 756
Db 1695 CTCAAACTCTCTGGGCTCAAGCGATCCTCCACCTTAGCCTCCCAAAGTACTGGGATTATA 1754
QY 757 ggtgtgagccacagtgccctggcctaattttcttctgtgatcaaaatccagggtttaatgtt 816
Db 1755 GGTGTGAGCCACAGTGCCTGGCCTAATATTATTTCTGTGATCAAAATTCAGGTTTAATGTT 1814
QY 817 ttgtgtaagaatttccctacggaattcgtgacttattttgtcatttagaggttcataaa 876
Db 1815 TTTGGTTAAGAAATTTCTACGTGAATTCGTGTACTTTATTTTGTCAATTAGAGTTCATAA 1874
QY 877 tattagggtttattttctaaatagaatgtttaaactaaataataacttcaaaacgctctag 936
Db 1875 TATTAGGGTTTATTCTTAAATAGATACTTTAAACTAAATATACTTCAAAACGCTCTAG 1934
QY 937 tttagtagctacccgtttgttggaatgaaattttctgtactgaaagaaacaaaagcct 996
Db 1935 TTTAGTAGTACCGTGTGTTGGATTGAAATTTCTGATACTGAAAAGAACAAAAGCCT 1994
QY 997 gcctttctgccc 1009
Db 1995 GCCTTTCTGCCCA 2007
RESULT 8
AR174329 AR174329 381 bp DNA linear PAT 17-DEC-2001
LOCUS Sequence 7 from patent US 6306653.
DEFINITION AR174329
ACCESSION AR174329
VERSION AR174329.1 GI:17914649
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 381)


```
AR174333/c
LOCUS       AR174333               311 bp    DNA    linear    PAT 17-DEC-2001
DEFINITION   Sequence 11 from patent US 6306653.
ACCESSION    AR174333
VERSION      AR174333.1  GI:17914653
KEYWORDS     Unknown.
SOURCE       Unknown.
ORGANISM     Unclassified.
REFERENCE    1 (bases 1 to 311)
AUTHORS      Papsidero, L.D., Dyster, L.M. and Frustaci, J.M.
TITLE        Detection and treatment of breast disease
JOURNAL      Patent: US 6306653-A 11 23-OCT-2001;
FEATURES     Location/Qualifiers
             1..311
             /organism="unknown"
BASE COUNT   65 a 77 c 95 t 2 others
ORIGIN
10.0%: Score 202; DB 6; Length 311;
Best Local Similarity 99.3%; Pred. No. 2.8e-99;
Matches 302; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 208 agccatactccctccctccagctgtgcagcgaggtttccacatcatatttccagaag 267
Db 311 AGCCATACTCCCATTCGCTCCAGCTGTGCACGGAGGTTTCACATCATATTTCCAGAAG 252
Qy 268 gctcctgaaagagtgatgtgtcgcacccagagagctgagtgaggattgtgactggc 327
Db 251 GCTCCTGAAAGAGTGAATATGTGTGCGCATCCAGAGAGCTGATGGGATTGTGACTTGGC 192
Qy 328 tgcgtcatccttcattcagcgacagagaatctgtgcagcccgacacaccatactgt 387
Db 191 TCGTGTCTATCTTCATGTCACGCCACAAANATCTGTGTGAGCCCGCACACCACTACTGT 132
Qy 388 taagcagtggaatgaagtgcagctgccaaagaaatgtttaaaggaaatgtttgcccacag 447
Db 131 TAAGCAGTGGATGAAGTGCAGAGCTGCCAANAAATGTTAAAGGAAATGTTTGCCACAG 72
Qy 448 gaagaacaccatggcgaagaggaaacagtaacaggggccacatcaggggaaacacgaaacata 507
Db 71 GAAGAAACACCATGGCAAGAGGAACAGATAACAGGGGCACATCAGGGGAAACACGAAACATA 12
Qy 508 cggc 511
Db 11 CGGC 8

RESULT 12
AR174330/c
LOCUS       AR174330               104 bp    DNA    linear    PAT 17-DEC-2001
DEFINITION   Sequence 8 from patent US 6306653.
ACCESSION    AR174330
VERSION      AR174330.1  GI:17914650
KEYWORDS     Unknown.
SOURCE       Unknown.
ORGANISM     Unclassified.
REFERENCE    1 (bases 1 to 104)
AUTHORS      Papsidero, L.D., Dyster, L.M. and Frustaci, J.M.
TITLE        Detection and treatment of breast disease
JOURNAL      Patent: US 6306653-A 8 23-OCT-2001;
FEATURES     Location/Qualifiers
             1..104
             /organism="unknown"
BASE COUNT   41 a 24 c 15 g 24 t
ORIGIN

Query Match 5.2%; Score 104; DB 6; Length 104;
Best Local Similarity 100.0%; Pred. No. 1.7e-45;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 745 actgggattatagtgtagccacagtgctggcctaattattttctgtgatcaaatc 804
Db 104 ACTGGGATTATAGTGTGAGCCACAGTGCTGCGCTTAATATTTCTTGTGATCAATTC 45
Qy 805 aggtttaatgtttttgttgaagaatttcctacgtgaattcgtgt 848
Db 44 AGGTTTAATGTTTTTGTGAAGAATTTCTTACGTGAATTCGTGT 1

RESULT 13
AC020929
LOCUS       AC020929               129427 bp    DNA    linear    HTG 23-APR-2001
DEFINITION   Homo sapiens chromosome 19 clone CTD-2223D2, WORKING DRAFT
SEQUENCE, 11 ordered pieces.
ACCESSION    AC020929
VERSION      AC020929.4  GI:13752692
KEYWORDS     HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE       human.
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 129427)
AUTHORS      DOE Joint Genome Institute.
TITLE        Sequencing of Human Chromosome 19
JOURNAL      Unpublished
REFERENCE    2 (bases 1 to 129427)
AUTHORS      DOE Joint Genome Institute.
TITLE        Direct Submission
JOURNAL      Submitted (12-JAN-2000) Production Sequencing Facility, DOE Joint
COMMENT      Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Apr 23, 2001 this sequence version replaced gi:9964758.
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov
-----
Project Information
Center Project Name: 713255, BC678279
Center clone name: CITB-H1_2223D2
-----
Summary Statistics
Consensus quality: 117979 bases at least Q40
Consensus quality: 124608 bases at least Q30
Consensus quality: 126982 bases at least Q20
Estimated insert size: 14240; agarose-fp estimation
Quality coverage: 6.99 in Q20 bases; agarose-fp estimation
Quality coverage: 7.75 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 11 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 11717: contig of 11717 bp in length
* 11718 11817: gap of unknown length
* 11818 34238: contig of 22421 bp in length
* 34239 34338: gap of unknown length
* 34339 48284: contig of 13946 bp in length
* 48285 48384: gap of unknown length
* 48385 80643: contig of 32259 bp in length
* 80644 80743: gap of unknown length
* 80743 88722: contig of 7979 bp in length
* 88722 91761: contig of 2939 bp in length
* 91762 91861: gap of unknown length
* 91862 99828: contig of 7967 bp in length
* 99829 109529: gap of unknown length
* 109529 109529: contig of 9601 bp in length
```

* 109530 109629: gap of unknown length
 * 109630 121814: contig of 12185 bp in length
 * 121815 121914: gap of unknown length
 * 121915 124932: contig of 3018 bp in length
 * 124933 125032: gap of unknown length
 * 125033 129427: contig of 4395 bp in length.

FEATURES

source

1. 129427
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="19"
 /clone="CTD-2223D2"
 /clone_lib="Caltech human BAC library D"
 31155 a 34031 c 31504 g 31737 t 1000 others

BASE COUNT 31155 a 34031 c 31504 g 31737 t 1000 others
 ORIGIN

Query Match 3.4%; Score 68; DB 2; Length 129427;
 Best Local Similarity 100.0%; Pred. No. 1.1e-25;
 Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1184 gaggctcactgtgcccagactgagtgagtcagtgacacatctcgctcattgcaacct 1243
 |||||
 Db 85637 GAGTCACACTCTGTCCAGGCTGGAGTGACGTGGCAACATCTCGCTCATTTGCAACCT 85696

QY 1244 ccgcctccc 1251
 |||||
 Db 85697 CCGCCTCC 85704

RESULT 14

AC021625

LOCUS

AC021625 Homo sapiens clone RP11-384E6, WORKING DRAFT SEQUENCE, 34 unordered
 pieces.

ACCESSION

AC021625.2

VERSION

HTG; HTGS_PHASE1; HTGS_DRAFT.

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 164652)

Birren,B., Linton,L., Nussbaum,C. and Lander,E.

Homo sapiens, clone RP11-384E6

Unpublished

2 (bases 1 to 164652)

Birren,B., Linton,L., Nussbaum,C., Lander,E., Abraham,H., Allen,N.,

Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,

Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,

Chospel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P.,

Dearellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,

Ferrelia,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,

Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,

Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,

Landers,T., Lehoczkv,J., Levine,R., Lieu,C., Liu,G., Locke,K.,

Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,

McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,

Norman,C.H., O'Connor,T., O'Donnell,P., Oliver,T.M., Peterson,K.,

Pierre,N., Pisanli,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,

Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,

Stojanovic,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,

Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,

Zimmer,A. and Zody,M.

Direct Submission

Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Mar 28, 2000 this sequence version replaced gi:6705474.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information

Center project name: L5880

Center clone name: 384_E_6

----- Summary Statistics

Sequencing vector: M13; M7815; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 145359 bases at least Q40
 Consensus quality: 155420 bases at least Q30
 Consensus quality: 159351 bases at least Q20
 Insert size: 161352; sum-of-contigs
 Quality coverage: 4.0 in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 34 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence.
 * as soon as it is available and the accession number will
 * be preserved.

1 1083: contig of 1083 bp in length
 * 1084 1183: gap of 100 bp
 * 1184 2479: contig of 1296 bp in length
 * 2480 2579: gap of 100 bp
 * 2580 4311: contig of 1732 bp in length
 * 4312 4411: gap of 100 bp
 * 4412 5951: contig of 1540 bp in length
 * 5952 6051: gap of 100 bp
 * 6052 7761: contig of 1710 bp in length
 * 7762 7861: gap of 100 bp
 * 7862 9630: contig of 1769 bp in length
 * 9631 9730: gap of 100 bp
 * 9731 11763: contig of 2033 bp in length
 * 11764 11863: gap of 100 bp
 * 11864 14294: contig of 2431 bp in length
 * 14295 14394: gap of 100 bp
 * 14395 16731: contig of 2337 bp in length
 * 16732 16831: gap of 100 bp
 * 16832 20279: contig of 3448 bp in length
 * 20280 20379: gap of 100 bp
 * 20380 23751: contig of 3372 bp in length
 * 23752 23851: gap of 100 bp
 * 23852 26903: contig of 3052 bp in length
 * 26904 27003: gap of 100 bp
 * 27004 31228: contig of 4225 bp in length
 * 31229 31328: gap of 100 bp
 * 31329 35336: contig of 4008 bp in length
 * 35337 35436: gap of 100 bp
 * 35437 38469: contig of 3033 bp in length
 * 38470 38569: gap of 100 bp
 * 38570 42643: contig of 4074 bp in length
 * 42644 42743: gap of 100 bp
 * 42744 47087: contig of 4344 bp in length
 * 47088 47187: gap of 100 bp
 * 47188 51475: contig of 4288 bp in length
 * 51476 51575: gap of 100 bp
 * 51576 56093: contig of 4518 bp in length
 * 56094 56193: gap of 100 bp
 * 56194 59387: contig of 3194 bp in length
 * 59388 59487: gap of 100 bp
 * 59488 63987: contig of 4500 bp in length
 * 63988 64087: gap of 100 bp
 * 64088 69347: contig of 5260 bp in length
 * 69348 69447: gap of 100 bp
 * 69448 74268: contig of 4821 bp in length
 * 74269 74368: gap of 100 bp
 * 74369 79365: contig of 4997 bp in length
 * 79366 79465: gap of 100 bp
 * 79466 83211: contig of 3746 bp in length

* 83212 83311: gap of 100 bp
 * 83312 88101: contig of 4790 bp in length
 * 88102 88201: gap of 100 bp
 * 88202 93785: contig of 5584 bp in length
 * 93786 93885: gap of 100 bp
 * 93886 100407: contig of 6522 bp in length
 * 100408 100507: gap of 100 bp
 * 100508 108158: contig of 7651 bp in length
 * 108159 108258: gap of 100 bp
 * 108259 115631: contig of 7373 bp in length
 * 115632 115731: gap of 100 bp
 * 115732 125488: contig of 9757 bp in length
 * 125489 125588: gap of 100 bp
 * 125589 136013: contig of 10425 bp in length
 * 136014 136113: gap of 100 bp
 * 136114 149110: contig of 12997 bp in length
 * 149111 149210: gap of 100 bp
 * 149211 164652: contig of 15442 bp in length.

FEATURES

Source
 1..164652
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="RPCI-11 Human Male BAC"
 1..1083
 /note="assembly_fragment"
 1184..2479
 /note="assembly_fragment"
 2580..4311
 /note="assembly_fragment"
 4412..5951
 /note="assembly_fragment"
 6052..7761
 /note="assembly_fragment"
 7862..9630
 /note="assembly_fragment"
 9731..11763
 /note="assembly_fragment"
 11864..14294
 /note="assembly_fragment"
 14395..16731
 /note="assembly_fragment"
 16832..20279
 /note="assembly_fragment"
 20380..23751
 /note="assembly_fragment"
 23852..26903
 /note="assembly_fragment"
 27004..31228
 /note="assembly_fragment"
 31329..35336
 /note="assembly_fragment"
 35437..38469
 /note="assembly_fragment"
 38570..42643
 /note="assembly_fragment"
 42744..47087
 /note="assembly_fragment"
 47188..51475
 /note="assembly_fragment"
 51576..56093
 /note="assembly_fragment"
 56194..59387
 /note="assembly_fragment"
 59488..63987
 /note="assembly_fragment"
 64088..69347
 /note="assembly_fragment"
 69448..74268
 /note="assembly_fragment"
 74369..79365
 /note="assembly_fragment"
 79466..83211

/note="assembly_fragment"
 83312..88101
 /note="assembly_fragment"
 88202..93785
 /note="assembly_fragment"
 93886..100407
 /note="assembly_fragment"
 clone_end:T7
 vector_side:right
 100508..108158
 /note="assembly_fragment"
 108259..115631
 /note="assembly_fragment"
 Query Match 3.4%; Score 68; DB 2: Length 164652;
 Best Local Similarity 100.0%; Pred. No. 1.le-25;
 Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1184 gagtctcactctgttcccgagctggagtcagtgccacaaatctcgctcattgcaacct 1243
 Db 14801 GAGTCTCAGTCTGTGTTCCCGAGCTGGAGTCAGTGGCACAAATCTCGCTCATTTGCCAACCT 14860
 QY 1244 ccgcctcc 1251
 Db 14861 CCGCCTCC 14868
 RESULT 15
 AC066608
 LOCUS
 DEFINITION Homo sapiens chromosome 3 clone RP11-785A7 map 3p, complete
 ACCESSION AC066608
 VERSION AC066608.5 GI:12745084
 KEYWORDS HTG.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS
 1 (bases 1 to 172805)
 Xiong,H., Zhou,Y., Dong,H., Lin,W., Chen,B., Zhang,C., Zhang,Y.,
 Cai,Z., Ying,H.F., Wang,H., Gu,W., Zhu,G., Tu,Y., Zhang,X., Jia,J.,
 Shen,H., Zhang,D., Wu,C., Lu,G., Zhong,M., Jiang,H., Ren,S., Fu,G.,
 Chen,Z. and Huang,M.
 TITLE
 Chromosome 3p genomic sequence
 JOURNAL
 REFERENCE
 AUTHORS
 2 (bases 1 to 172805)
 Zhang,X., Hu,S., Dong,W., Wang,J., Zhang,Y., Zhang,H., Liu,B.,
 Bao,W., Sun,Y., Wu,Q., Wang,H., Yang,X., Cheng,C., Wang,Y., Niu,Y.,
 Qi,X., Li,T., Zhang,H., Liu,N., Wu,D., Yu,B., Fan,H., Liu,Y.,
 Li,G., Li,C., Bao,Q., Bao,J., Wang,X., Song,L., Zhang,L., Guo,D.,
 Huang,F., Zhang,G., Li,J., Bian,X., Zhang,M., Li,L., Feng,X., Yu,J.
 and Yang,H.
 TITLE
 Direct Submission
 JOURNAL
 REFERENCE
 AUTHORS
 Submitted (25-APR-2000) Human Genomic Center, Institute of
 Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
 100101, P.R.China
 3 (bases 1 to 172805)
 Bao,J., Bao,Q., Bao,W., Bian,X., Cao,T., Chen,C., Chen,J., Ding,H.,
 Dong,W., Fan,H., Feng,X., Gong,J., Guan,Q., Gu,X., Guo,D., Guo,Z.,
 He,L., Hu,S., Huang,F., Jin,Y., Kang,N., Li,C., Li,C., Li,F.,
 Li,G., Li,J., Li,L., Li,S., Li,T., Liu,Y., Liu,N., Liu,B., Liu,Y.,
 Li,W., Li,W., Li,Y., Luo,C., Luo,J., Niu,Y., Qi,O., Qi,X., Song,L.,
 Song,S., Sun,M., Sun,W., Sun,Y., Tan,X., Tao,R., Wang,H., Wang,J.,
 Wang,J., Wang,J., Wang,L., Wang,L., Wang,R., Wang,X., Wang,X.,
 Wang,Y., Wu,D., Wu,Q., Xie,F., Xuan,Z., Xue,Y., Yan,C., Yang,X.,
 Yu,B., Zeng,Y., Zhang,G., Zhang,H., Zhang,H., Zhang,L., Zhang,M.,
 Zhang,X., Zhang,X., Zhang,Y., Zhang,Y., Zhang,Z., Zhu,B., Zhu,N.,
 Yu,J. and Yang,H.
 TITLE
 Direct Submission
 JOURNAL
 REFERENCE
 AUTHORS
 Submitted (11-FEB-2001) Human Genomic Center, Institute of
 Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
 100101, P.R.China

```
COMMENT      On Feb 11, 2001 this sequence version replaced gi:8101281.
-----Genome Center
Center:Beijing Center
Center code:Beijing
Website:http://hgc.igtp.ac.cn
http://www.genomics.org.cn
Contact:hgc@igtp.ac.cn
-----Project Information
Center project name:1% project
Center clone name: RP11-785A7
-----Summary Statistics
Sequencing vector: pUC18; 100% of reads
Chemistry: Dye-terminator: ET 55% of reads
Chemistry: Dye-terminator Big Dye; 45% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 176043 bases at least Q40
Consensus quality: 176221 bases at least Q30
Consensus quality: 176261 bases at least Q20
Insert size: 172805; sum-of-contigs
Quality coverage: 8.12x in Q20 bases;sum-of-contigs
-----
FEATURES             Location/Qualifiers
     source            1..172805
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /chromosome="3"
                        /map="3p"
                        /clone="RP11-785A7"
BASE COUNT      53766 a 33397 c 32927 g 52715 t
ORIGIN

Query Match      3.3%; Score 67; DB 9; Length 172805;
Best Local Similarity 100.0%; Pred. No. 3.7e-25;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1169 ttttttttgagacagagtctcactctgtgtgccagctgagtgagtgacacaaatctc 1228
      |||||
Db 76931 TTTTGTGAGACAGAGAGTCTCACTCTGTGTCGCCAGGCTGGAGTGCAGTGCACAAATCTC 76890

Qy 1229 ggctcat 1235
      |||||
Db 76891 GGCTCAT 76897

Search completed: October 7, 2002, 13:05:27
Job time: 9417 sec
```

THIS PAGE BLANK (USPTO)